Face and object processing: What changes with experience?

Rachel Robbins
B.Sc. (Hons.), *ANU*

Submitted for the degree of Doctor of Philosophy

School of Psychology, The Australian National University

January 2005

I certify that, aside from the usual support and advice provided by my supervisor Dr Elinor McKone, the ideas and research in this thesis are solely my own work. To the best of my knowledge there is no material in this thesis published or written by another person except where appropriately acknowledged.

Examiners should note that Experiment 1 (Chapter 3) was submitted and examined as my Honours thesis (2000) and is therefore not eligible for examination for the degree of Doctor of Philosophy. No other work in this thesis has been submitted for examination for a degree at any university.

Rachel Robbins

Rachel Robbins

Previously published material appearing in this thesis.

Robbins & McKone (2003), Can holistic processing be learned for inverted faces? *Cognition, 88*(1), 79-107. **Experiments 1-3.**

Rhodes, G., Robbins, R., Jaquet, E., McKone, E., Jeffery, L., & Clifford, C. (in press). Adaptation and Face Perception – How Aftereffects Implicate Norm Based Coding of Faces. In G. Rhodes & C. Clifford (Eds.), *Fitting the Mind to the World: Adaptation and aftereffects in high level vision.* Oxford: Oxford University Press. **Includes a sub-set of the data from Experiment 10, attributed in the text as Robbins and McKone (2004), unpublished raw data.**

Acknowledgements

First of all I thank my supervisor Dr Elinor McKone. She has been extremely helpful at every stage of the PhD process. I have particularly appreciated her insistence that I keep growing towards her very high standards. Thanks also to my advisors: to Dr Mark Edwards for his help and suggestions, particularly with regard to neural models (Chapter 6); and Dr Cobie Brinkman for helpful advice on primates.

Thank you to all the people who participated in my experiments, who gave up hours of their time in the name of science (you know who you are)! Particular thanks to the twins who let me take hundreds of photographs of them for Experiments 1-3, and the dog experts who participated in Experiments 7-9. Thanks also to Mary Conheady of the Victoria Labrador Retrievers Association for her assistance in finding pictures of dogs.

Thanks to the PARSA people, especially Di, James, Louise, Fiona, Brett, Patrice and Sandra. I appreciate all the different perspectives I got on uni life though my PARSA work, plus the friendship and support you've all given. Thanks also to the other psych PhD students, especially Nic C who shared an office with me!

Special thanks to all the people who helped keep me sane and put up with me at my worst. In no particular order, thanks to my wonderful partner Benedict, my brother bug and his gf Nic, my parents, Nine, my house-mates, and all my friends (I'm not going to try and name everyone because I know I will forget someone and make them sad :). I love you all!

This thesis was fuelled by anime music, Belgian chocolate & cinnamon tic tacs.

Abstract

At the most general level, this thesis examines the ways in which face and object processing change with experience, and the ways in which they do not. Two theoretical topics are of particular interest. The first is the origin of holistic/configural processing: under normal circumstances, this is limited to faces, seen upright; here I explore whether it can develop with experience for inverted faces or for objects-of-expertise (thus testing the expertise hypothesis). The second is related to understanding face-space, and the norm-based coding of individual face identity. Timescales of experience considered range from 2 mins of adaptation to a distorted face, to 20 years at judging show-dogs.

Empirical chapters are organised as four more-or-less independent papers, each addressing somewhat distinct theoretical questions. The first empirical chapter tests whether extensive training with inverted faces is sufficient to change processing style to that used for upright faces (i.e. holistic). The second empirical chapter examines which tests are best for showing differences between faces and objects, in subjects with no particular expertise with the target object class (i.e., dog novices). The third empirical chapter tests whether dog-show judges (i.e., dog experts) process dogs in the same way as they do faces. The fourth empirical chapter tests whether adaptation to a distorted face changes perception in relation to a norm, whether any such changes are based on previous experience with faces, and whether adaptation transfers between upright and inverted faces when a relational (spacing) distortion is used.

Taken together, results suggest that upright-face-like processing does not occur for inverted faces or dogs after hours to years of practice; that previous experience has a strong effect on adaptation; that adaptation for both upright and inverted faces occurs with respect to a norm, but that the adaptation for upright and inverted faces occurs in different populations of neurons. In terms of methodology, results suggest that some tasks are better than others for asking questions about the domain-specificity of faces, and that adaptation to simple face-shape distortions is a good way to investigate the structure of "face-space". Overall, the thesis concludes that there is little or no evidence to support the idea that practice changes processing of non-face objects or inverted faces in such a way that it becomes like that for upright faces. Possible origins of domain specificity (based on previous literature) are also discussed.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

CHAPTER 1: HOW ARE FACES "SPECIAL"?

1.1 Overview of the thesis

The general intent of this thesis is to examine the role and effects of experience on visual face and object processing. In particular I am interested in how different amounts of experience may be related to holistic or configural processing, namely the style of cognitive processing often argued to be "special" to faces, and considered to be an important part of what supports humans impressive ability to distinguish between many individual faces. The different amounts of experience to be examined range from minutes to many years. As an example of the former, I investigate changes in what appears normal after a two minute adaptation to distorted faces and how this relates to face-representations derived from long-term experience. As an example of the latter, I investigate how dog experts with 20 years experience process pictures of dogs.

One aim of this thesis is to contribute to the debate on whether holistic/configural processing is domain specific, that is, limited to faces, and indeed, to upright faces. As a review of previous empirical studies will show, upright faces are special at least compared to many other stimuli. The question is whether holistic/configural algorithms can be applied to non-face objects, or inverted (upside-down) faces under certain circumstances. I am primarily interested in the theory that within-class discrimination plus experience produces holistic processing; here, the amount of practice required might be anything from hours to years. This expertise hypothesis is contrasted with the theory that holistic/configural processing is genuinely domain-specific, and thus only upright faces are holistically processed. (Were such domain-specific processing supported, it might have some innate basis; however, this possibility is not empirically tested in this thesis.) These theories will be fully explored in the literature review and experimental chapters.

Most of the empirical work in the thesis addresses this first aim of evaluating domain specificity. The first experimental chapter addresses the question of whether training in an experimental setting is sufficient to develop upright-face-like processing for inverted faces. Inverted faces are an obvious comparison stimulus to upright faces in that they have the same low-level properties (e.g., spatial frequency, brightness, etc.) but, without practice, are not processed holistically. The chapter explores the development, or otherwise, of holistic processing with 10 hours of exposure to inverted faces, an amount of practice which has been suggested to induce holistic processing for the artificial objects "greebles", and which is well in excess of the number of trials required to remove inversion effects for visual recognition of objects.

Subsequent work then explores the emergence, or otherwise, of holistic processing with much greater levels of experience. In this case, the stimulus class considered is a type of object – labrador dogs – rather than inverted faces, and the subjects of interest are real-world experts with many years of experience in individual identity level judgements for this class (e.g., labrador dog-show judges). I argue that labradors provide a good comparison for face stimuli in several ways, the most important being that they are living things with genetic diversity and thus, as for faces, individuals differ from one another on a great many dimensions; this is not the case for man-made objects such as greebles. The chapter reports tests for face-like processing of labradors on several tasks, some relevant to holistic processing (inversion effects, composite task) and one that tests for a non-holistic phenomenon (sensitivity to reversed contrast), in labrador experts. Dogs have not previously been compared to faces on two of these tests even in novices. Thus, the chapter before the one on dog experts presents results on the tasks for subjects who have no particular expertise with dogs beyond general familiarity and knowledge of a canonical upright. The results for these "dog novices" confirm that within-class discrimination alone is not sufficient to induce face-like processing on any of the measures. The primary aim of this study with novices is to consider which tests of face recognition are best for testing the expertise hypothesis, by establishing which show effects for faces and not objects (a qualitative difference) rather than having only a disproportionately large effect for faces compared to objects (a quantitative difference).

The last study relevant to the domain specificity hypothesis returns to a comparison of upright and inverted faces, using a method that explores the effect of very short-term experience – a two minute adaptation period – on the perception of facial distortion. The result of interest concerns the transfer of adaptation across orientations. There is substantial behavioural evidence that upright and inverted faces are processed differently at a cognitive level. However, the area of the brain most strongly activated by faces (compared to other objects) is also activated by inverted faces. A single study has used adaptation to show that upright and inverted faces are processed by different neural populations. Lack of transfer of adaptation between upright and inverted faces would provide another test of this, and show that different neural populations are used for the particular distortion I have tested.

The second aim of this thesis is to use the effects of experience to understand the representations of upright faces as individuals. This can be done using temporary face-adaptation in the theoretical context of a "face-space". As noted above, holistic processing assists us in recognising individuals. Each individual face may also be thought of as belonging to some kind of multi-dimensional face-space, presumed to be built up of faces seen over the life of an individual. Of particular interest is the form that adaptation takes, whether it is related to a norm in face-space, and how it can be modelled in neural terms. I also examine whether some dimensions in face-space are more adaptable than others and whether this is related to differences in the natural population of faces experienced in everyday life. Thus, this part of the thesis deals with how short-term exposure may change face representations built up over longer-term experience.

Several aspects of the literature on face processing form the background to the experimental studies. These are reviewed, along with the basic theories relevant to the experiments, in Chapters 1 and 2 of the thesis. Chapter 3 then contains the experiments on processing of inverted faces after extensive training (Experiments 1-3). Chapter 4 contains experiments comparing results for faces versus labradors on several tests in non-dog-experts (Experiments 4-6), and Chapter 5 contains the corresponding experiments in dog experts (Experiments 7-9). Chapter 6 presents the experiments on adaptation for different dimensions in face-space (Experiments 10 & 11), that on transfer of adaptation between upright and inverted stimuli (Experiment 12) and some preliminary neural models of the results. Chapter 7 is the General Discussion, which draws together the results of all experiments, and considers how these relate to each other and to the previous literature.

1.2 Review of empirical evidence on whether faces are special.

This section will review literature relevant to the question of whether upright and inverted faces are processed differently, and whether faces are processed differently from other objects. Literature from behavioural tasks as well as relevant work from neuroscience will be reviewed; note that all experiments reported in the present thesis are behavioural experiments on adult human subjects. Also note that in this thesis "face processing" means individual identity not, for example, emotion or attractiveness; previous research has shown that identity is separable from these other aspects of face processing (e.g., Haxby, Hoffman, & Gobbini, 2000; Schweich & Bruyer, 1993).

In this chapter only object processing in non-experts will be discussed. That is, the subjects of the reviewed experiments will be ordinary people with a general familiarity for the class of objects tested, rather than extensive practice at discriminating similar individuals of that class. Note that all people are generally agreed to be expert at face recognition, unless they have some kind of disorder which prevents them from recognising people normally (e.g., prosopagnosia, or possibly autism).

1.2.1 The disproportionate inversion effect on memory for faces.

In behavioural experiments, early studies on memory for upright and inverted stimuli suggested that faces were processed differently from other objects. On a task of recognition memory, faces were more adversely affected by inversion than other stimuli which also have a canonical upright (e.g., aeroplanes, houses, or stick figures displaying different actions; Yin, 1969). This disproportionate inversion effect for faces occurred even when the non-face stimuli (pictures of period costumes) were easier to recognise than faces in the upright orientation (Yin, 1969). The effect has also been replicated for other stimuli including pictures of buildings (Scapinello & Yarmey, 1970) and landscapes (Diamond & Carey, 1986), dogs (in non-experts; Diamond & Carey, 1986) and dog faces (Scapinello & Yarmey, 1970).

Yin interpreted his results as indicating that upright faces are processed in a holistic or configural manner whereas inverted faces and objects are processed in a part-based manner. However, Valentine (1988) noted that a disproportionate inversion effect does not provide any direct evidence of holistic processing, and moreover, that the disproportionate inversion is only a quantitative not qualitative difference and does not mean that faces are special *per se*. Before I describe more recent evidence arguing directly that upright faces are holistically processed, I use the next section to describe what is and is not meant by holistic processing in the present thesis.

1.2.2 Exactly what is holistic processing for faces?

In the face recognition literature, the terms "holistic" and "configural" processing are taken to mean some kind of extraction or integration of information across the entire face region (excluding hair). The exact definition of these terms remains a matter of debate (e.g., see Maurer, Le Grand, & Mondloch, 2002, for a discussion). They vary from meaning processing information from the whole face with no decomposition into local parts (e.g., Farah, 1996; Moscovitch, Winocur, & Behrmann, 1997; Tanaka & Farah, 1993) to initial part-decomposition followed by encoding relationships and/or spacing between multiple features (e.g., Rhodes, 1988). Note that faces are generally agreed to share a first-order configuration; that is, the same parts in the same basic order (for faces, two eyes above a nose above a mouth). Differences in spacing between features are then sometimes referred to as differences in the second-order relationships (Diamond & Carey, 1986). It is not clear exactly how the spacing between multiple sets of features is related to "holistic" processing (in the sense of processing an undifferentiated whole) except that changes in spacing between features will change the whole. Most authors agree that tests of these kinds of changes all access aspects of the same underlying process. In the present thesis, I use the term holistic to cover all versions of holistic or configural processing proposed (e.g., see Maurer et al., 2002) and will not distinguish between them. Holistic/configural processes are contrasted with local part or feature information (e.g., exact shape of eyes or lips).

It is also important to briefly note what is not meant by holistic processing in the context of this thesis. Many stimuli show some advantage for parts being processed in context. For example, letters are recognised more easily in the context of words than in isolation, or in jumbled letter strings (the word superiority effect, Reicher, 1969; Wheeler, 1970). Although this may be one kind of holistic processing it is not the kind used for recognising faces. In support of this, there exist patients who, although unable to recognise faces, can read normally (Humphreys & Rumiati, 1998; Moscovitch et al., 1997). Similarly, the ability to use closure to recognise Mooney faces or objects (very high contrast pictures with no grey tones) is not the same as holistic processing of faces, although Mooney faces/objects are easier to see when you have been told what to look for. This claim is supported by the finding that a patient with impaired object recognition but intact face recognition can accurately describe Mooney faces (e.g., a young girl looking right) but not Mooney objects (Moscovitch, et al., 1997). Further, a double-dissociation exists with a patient unable to recognise faces, but able to recognise objects with missing segments or presented in visual noise (Duchaine, 2000). These findings suggest that closure is an earlier stage of visual processing to the holistic processing involved in face recognition. One final example of what is not meant by holistic processing in this thesis must be given. Gauthier, Curran, Curby, & Collins (2003) define holistic processing as "obligatory processing of all features of an object, even when subjects are instructed to attend selectively to one feature while ignoring others" (p.428). This definition as the inability to ignore notionally irrelevant stimuli would mean that anything with high automaticity of processing would be holistically processed with whatever else was presented at the same time. As a concrete example, this definition would suggest that words plus colours are holistically processed in the Stroop effect (Stroop, 1935). Such an attention-based definition is not what is usually meant by holistic processing in the literature on faces. This thesis will examine holistic processing only in the strict sense used in the face literature.

1.2.3 Evidence of holistic processing for upright but not inverted faces.

Subsequent to Valentine's (1988) review, several paradigms have directly shown the special nature of upright face processing. In Young, Hellawell, and Hay's (1987) composite effect, the top half of one famous person's face was combined with the bottom half of a different famous person's face (e.g., Tony Blair's forehead with George Bush's chin). An example is given in Figure 1.1. For upright faces, when these two halves were physically aligned, subjects were slower to name either half (e.g., the top half) than when the two halves were offset. Given that simple response competition from the two halves is the same in both aligned and unaligned conditions, this indicates that perceptual fusing of the two halves (i.e. holistic processing) occurred when the halves were aligned. For inverted faces, in contrast, there was no difference in naming times for aligned and unaligned stimuli, indicating no holistic processing. This result has been replicated with both familiar and unfamiliar faces (Carey & Diamond, 1994; Hole, 1994; Le Grand, Mondloch, Maurer, & Brent, 2004). For unfamiliar faces, a matching version of the task is used (the variations of the task will be discussed in Chapter 5).

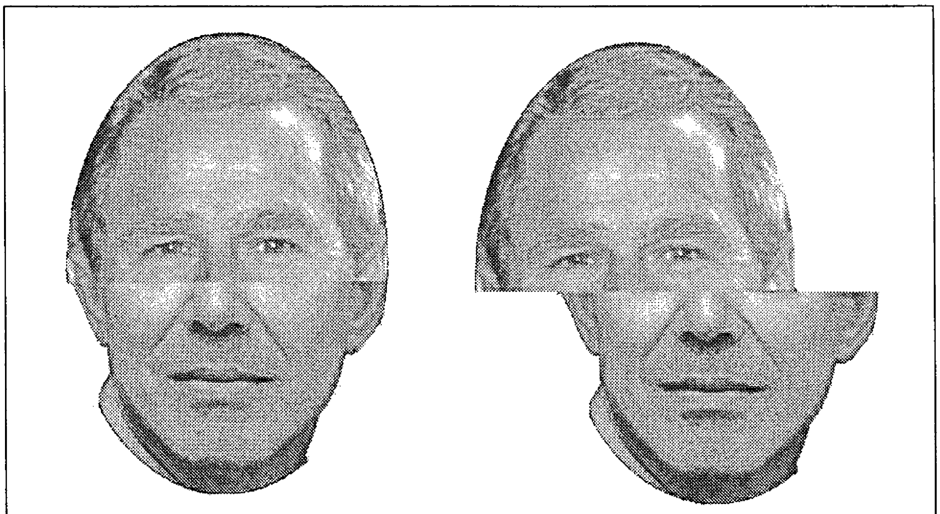


Figure 1.1. An example of stimuli used in the composite task, made up of the top-half of George Bush's face and the bottom half of Tony Blair's face. In Young et al. (1987) reaction times to name the top half were significantly slower in the aligned than the unaligned condition for upright stimuli, but there was no difference between conditions for inverted stimuli.

In Tanaka and Farah's (1993) part-whole paradigm, subjects first learned whole faces (e.g., Bill). In a subsequent memory test stimulus pairs were presented, either as isolated parts (e.g., Bill's mouth vs. Jim's mouth) with subjects asked to choose a specified part (e.g., Bill's mouth), or in the context of the whole face (e.g., Bill's mouth in Bill's face vs. Bill's mouth in Jim's face) with subjects asked to choose a specified face (e.g., Bill's face). An example is shown in Figure 1.2. In the upright orientation, memory for the face part was better in the whole face condition than in the isolated part condition, indicating strong integration of parts into wholes. This did not occur, however, for inverted faces or scrambled faces. Scrambled faces contain all the same face parts as normal faces but the first-order configuration has been changed (e.g., the mouth at the top, one eye and the nose in the middle, and the other eye at the bottom). Results for scrambled faces thus show that it is not only the presence of the parts in the correct upright orientation, but also the arrangement of those parts which is important. Variations on this paradigm include: an immediate memory version, where subjects are presented with a face and then presented with either two parts alone or two parts in the face (Davidoff & Donnelly, 1990; Donnelly & Davidoff, 1999)[1]; and a version where

[1] Note that the effect is referred to by Davidoff and Donnelly as a Complete Probe Advantage or CPA.

subjects are asked whether it is Bill's mouth in both part and whole conditions (rather than Bill's face in the whole condition; Tanaka & Sengco, 1997).

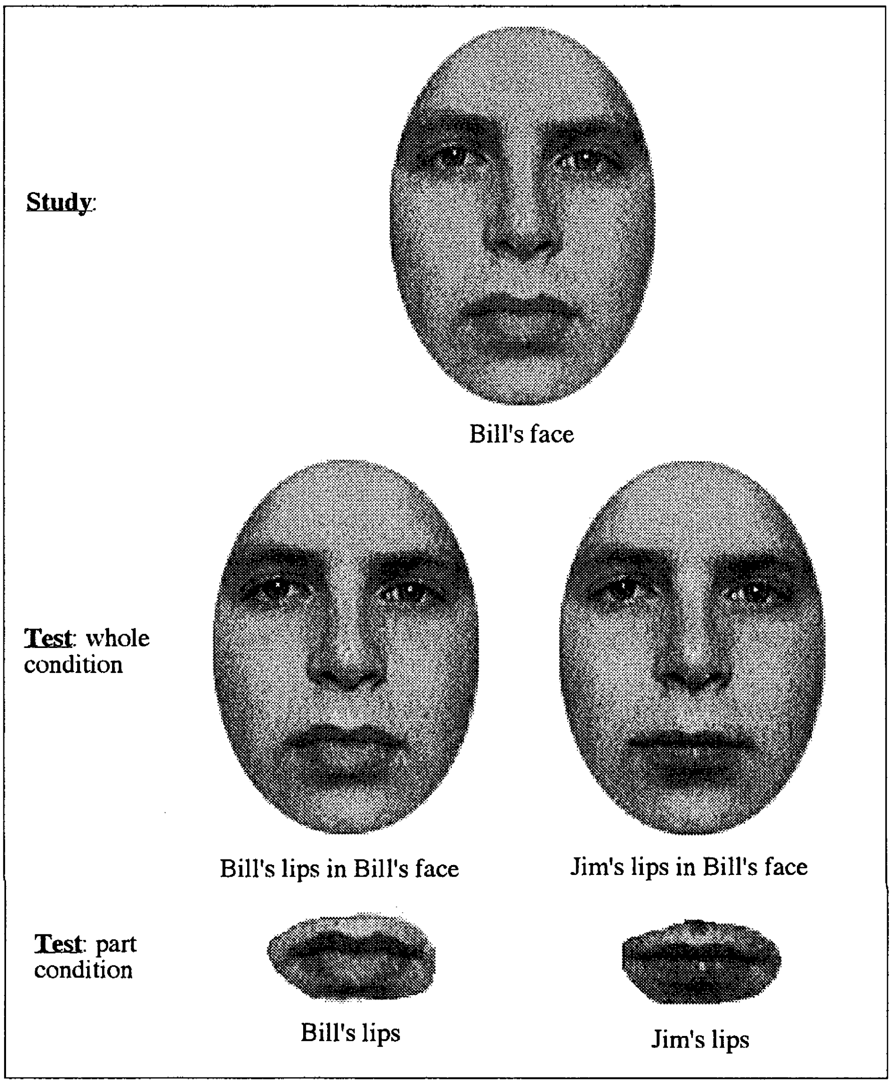


Figure 1.2. An example of stimuli used in the part versus whole task. Subjects are more accurate at remembering Bill's lips in the context of Bill's studied face (whole condition) than when presented alone (part condition). This difference occurs when all stimuli are upright, but not when they are inverted.

A range of techniques have been reported which fall under the heading of relational (or spacing) versus local feature alterations. The first of these was the Thatcher illusion reported by Thompson (1980), in which the eyes and mouth of a face (in the original case, Margaret Thatcher) are inverted while the rest of the face is unaltered. The resulting stimulus appears extremely grotesque when upright, but becomes only slightly odd when inverted. Murray, Yong, & Rhodes (2000) compared Thatcherised faces, and faces with altered spacing/relationship between features (e.g., the mouth moved abnormally far down the face), to faces where an individual local feature had been altered (in this case blacking out teeth or whiting out eyes). Ratings of "bizarreness" were taken at various orientations between upright and inverted and it was found that while ratings for normal and featurally altered faces decreased linearly with inversion, the change in ratings with rotation was more bell-shaped for spacing-altered and Thatcherised faces. Murray et al. conclude that spacing type changes (ones that alter the holistic nature of the face) are noticed well in the upright orientation, but poorly when the face is inverted. Changes to local features, conversely, are obvious in upright and inverted orientations (as well as intermediate ones). Other experiments have come to similar conclusions (e.g., Bartlett & Searcy, 1993; Gilchrist & McKone, 2003; Le Grand, Mondloch, Maurer, & Brent, 2001; Leder & Bruce, 1998; Rhodes, Brake, & Atkinson, 1993). An example of a spacing change, a featural change (of the kind used by Leder & Bruce, 1998) and a Thatcherised face are shown in Figure 1.3.

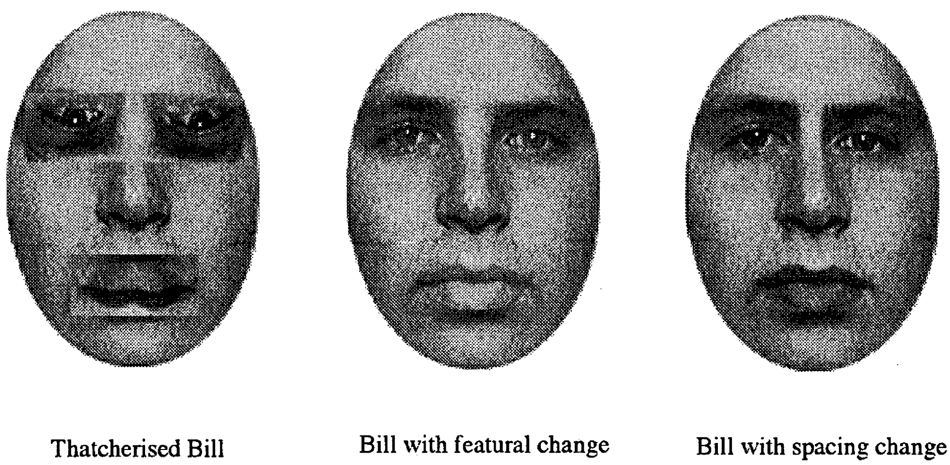


Figure 1.3. An example of a Thatcherised face (the fact that it looks less odd when inverted can be seen by turning the page upside-down), a featural change to the colour of mouth and eyes (of the kind used by Leder and Bruce, 1998), and a spacing/relational change in the distance between eyes. Original Bill is shown in Figure 1.2.

Two further techniques have been shown to isolate the configural/holistic component of face processing, and to show that it occurs only for upright faces. In the first, McKone, Martini & Nakayama's (2001) categorical perception in noise technique, subjects were trained to classify two faces and a continuum of morphs between them as either “Face 1” or “Face 2”. These were presented in heavy visual noise (e.g., random dots) to make recognition via local features unreliable. After up to 10 sessions of this classification practice, subjects were asked to make similarity judgements between pairs of faces, or to say which of two sequentially presented faces was “most like Face 1”. Categorical perception is defined as better discrimination around a perceived boundary between the two faces (e.g., for 40% and 60% Face 1 morphs if the boundary was 50%) than for morphs an equal distance apart but on the same side of the perceived boundary (e.g, 20% and 40% Face 1 morphs). McKone et al. found categorical perception for upright faces, but not for inverted faces or for isolated parts (noses).

In the second technique isolating configural processing, McKone's (2004) peripheral identification technique, faces were flashed at various distances from fixation. For both upright and inverted faces, identification accuracy reduced as horizontal distance from fixation increased, consistent with the degradation of low level visual processing that occurs in the periphery. The important finding was that the falloff with distance from fixation was faster for inverted faces than for upright faces. This produced a peripheral inversion effect, namely a reemergence of an inversion decrement in the periphery when subjects had previously been trained to identify upright and inverted faces to ceiling accuracy with central presentation. McKone (2004) argued that these results arose because (a) upright faces are processed in a holistic plus part-based manner while inverted faces are processed as parts only, (b) holistic processing integrates information from across larger regions of the face than does part-based processing, and (c) peripheral presentation degrades the information available from any single local region of the face (e.g., the nose alone, or a patch of skin on the left cheek). Together, these factors allow identification of upright faces to survive further into the periphery than identification of inverted faces. In direct support of this interpretation, no peripheral inversion effect was found for an isolated face part, namely the nose alone.

From the studies reviewed above, it is apparent that upright and inverted faces are processed qualitatively differently. Moreover the results for upright faces indicate that they are holistically/configurally processed, whereas inverted faces and scrambled faces are processed as a mostly unrelated collection of parts.

1.2.4 Results for objects using the above holistic processing paradigms.

I now review findings for the paradigms described above with objects as stimuli (note again that this is only in non-experts). For many of these paradigms there have been few (or, in the case of McKone et al.'s, 2001, categorical perception in noise task, no) objects tested, so for these cases the claim that the paradigms show that faces are special relies only on the distinctions between upright and inverted faces.

The composite effect has only been tested with one class of non-face objects, the novel objects known as "greebles". Greebles are vaguely animal-like rendered shapes (i.e., they look three-dimensional) with cylindrical "bodies", "heads" and four protrusions (three from the "head", one from the "body"; see Figure 1.4). Like faces, all greebles share a first-order configuration. In assessing the composite effect, Gauthier & Tarr (2002) required subjects to match the target half of aligned and unaligned greebles in a sequential same/different task. Aligned trials were matched approximately 42 ms faster than unaligned trials. Holistic processing would be indicated by aligned trials being matched slower than unaligned trails; thus greebles were not holistically processed.

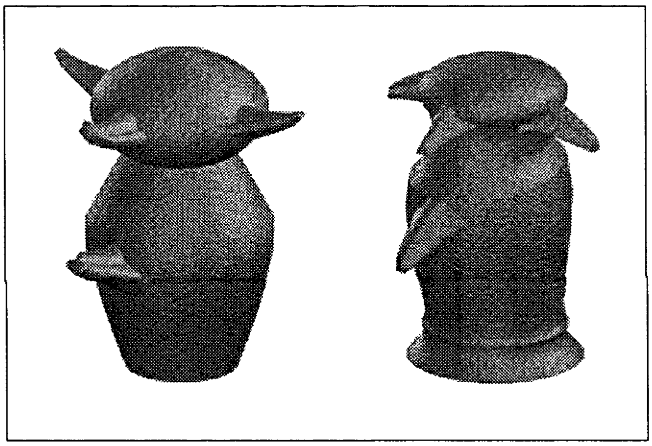


Figure 1.4. Example greebles courtesy of Michael J. Tarr (Brown University).

The part-whole paradigm has been tested on more object classes than any other test of holistic processing. This includes the long-term and short-term memory versions of the task, as well as a "transformed" version in which the part is tested in the whole but with a spacing change also made (e.g., for faces this might be Bill's mouth in Jim's face with Jim's eyes moved apart; Tanaka & Sengco, 1997). In the latter case the result for faces is that performance for the part in a transformed whole face is intermediate to performance for an isolated part or the part in the original whole face (in fact it is significantly better than performance for parts but significantly worse than performance for original whole faces). Using Tanaka and Farah's (1993) long-term memory version of the part-whole paradigm, no part-whole effect or transformed part-whole effect was obtained for houses (Donnelly & Davidoff, 1999; Tanaka & Farah, 1993; Tanaka & Sengco, 1997), and no part-whole effect was obtained for dog faces (Tanaka et al., 1996, cited in Tanaka & Gauthier, 1997). A small part-whole effect was found for car fronts and biological cells (approximately 8-16%; Tanaka et al., 1996, cited in Tanaka & Gauthier, 1997), but this was much smaller than for faces (approximately 25% in this study). For the artificial stimuli "greebles", small effects have also sometimes been found. There was no reliable effect in Gauthier, Williams, Tarr, and Tanaka (1998), the only difference being because one isolated part was below chance. There was a 6% part-whole effect in Gauthier and Tarr (1997); and a part-whole effect in d' (a measure of sensitivity) of 1.4, and a transformed part-whole effect in d' of approximately 0.5. Note that neither of the greeble studies included faces as a comparison group.

Using the immediate memory version of the paradigm (CPA), a small part-whole effect was found for houses (approximately 5-10%, Donnelly & Davidoff, 1999), and chairs (5% compared to 11% for faces, Davidoff & Donnelly, 1990). Thus, the part-whole effect may be like the inversion effect in that the effect is much larger for faces than for other objects, although a small effect does occur for objects.

Only one study has examined the Thatcher illusion with an object class, and then only on children. Rouse, Donnelly, Hadwin, and Brown (2004) asked children to identify which of two houses was "odd". To make Thatcherised houses the bay window and door were turned upside-down. This gave the impression of an inward curving, rather than an outward curving bay window. Rouse et al. compared the differences in accuracy to pick the odd stimulus between upright and inverted orientation for faces and houses. The inversion effect was 34% for faces, and 14% for houses (although note that the upright performance for houses was worse, making comparison of the size of the inversion effects difficult).

Sensitivity to spacing-based changes has not been tested with any natural object class. Many studies have changed the distance between parts of made-up objects, and some have even compared these to changes to shape of parts (e.g., Keane, Hayward, & Burke, 2003, Experiment 4). However, these were stimuli with no canonical upright, making comparison with tests using faces difficult (because for faces the finding is that both spacing and feature changes are noticed well upright, but that feature changes are noticed much better than spacing changes inverted). Obviously much more work needs to be done before conclusions can be drawn about whether there are differences between faces and objects on this paradigm.

Turning to the paradigms which isolate holistic processing for faces, the categorical perception in noise paradigm has not been used. The peripheral identification task has been tested for dachshund dogs. MacPherson (2001) showed that within-class discrimination of dachshunds did not produce the inversion effect in the periphery which is characteristic of holistic processing for faces. If anything, processing of inverted dogs was slightly better than processing of upright dogs, and both declined at the same rate as eccentricity increased.

In summary, while testing of more object classes would be valuable, it is clear that (in non-experts) objects do not receive the same type of holistic processing as faces. Objects show small inversion effects and small part-whole effects, but these are not nearly as big as for faces. Further, the composite and peripheral identification paradigms show no effect, at least for the limited classes of objects tested so far (interestingly this suggests that these paradigms may provide purer measures of face-like holistic processing that the other tasks).

1.2.5 Non-holistic behavioural paradigms on which faces and objects differ.

Contrast reversal. Effects of contrast reversal (reversing the luminance values of pixels so that the picture looks like a photographic negative) have been evaluated as potentially "special" to faces (e.g., Gauthier, et al., 1998). Many studies have shown that contrast reversal substantially impairs face recognition (e.g, Bruce & Langton, 1994; Johnston, Hill, & Carman, 1992; Kemp, Pike, White, & Musselman, 1996) and even impairs discrimination of spacing changes in faces (Kemp, McManus, & Pigott, 1990). This does not necessarily imply that contrast reversal only or primarily affects holistic processing for faces. Instead, it seems that contrast reversal disrupts information about shape-from-shading at an independent stage in the visual processing stream. Kemp, et al. (1990) showed that, in a match-to-sample task, inversion (turning the face upside-down) and contrast reversal each reduced accuracy in an additive fashion (i.e., there was no interaction between the two), indicating that inversion and contrast reversal affect face-recognition in different ways (also see Bruce & Langton, 1994). Confirming this result, Hole, George, and Dunsmore (1999) found that there was a composite effect (i.e., holistic processing) that was as large for contrast reversed faces as for original faces.

For objects, three studies have investigated contrast reversal effects. In the one published study, Gauthier et al. (1998) found no effect of contrast reversal for identifying previously learned greebles. In a conference presentation, Subramaniam & Biederman (1997) reported the same null result with a sequential same-different task for chairs. However, in both these studies the objects had strong part boundaries that are relatively insensitive to contrast reversal and thus processing may have been less reliant on shape-from-shading cues than is the case for faces. Reporting in a second conference presentation, Nederhouser, Mangini, Biederman, & Kazunori (2002) addressed this to some extent by using "blobs" (pictures of smooth three-dimensional objects with a number of random looking protrusions), and again found no effect of contrast reversal. Thus, results to date suggest that objects do not produce face-like contrast reversal effects.

Attention. Although attention is not the focus of this thesis, it is worth noting that upright faces may also be special with regard to attentional resources. For example, even though faces do not "pop-out" in visual search (e.g., Brown, Huey, & Findlay, 1997), patients with unilateral neglect can attend to the neglected side if something "makes a face" across the mid-line (e.g., Vuilleumier, 2000; Vuilleumier & Sagiv, 2001). That is, when two crosses or two digits were presented with a curved line below, all inside a circle (i.e., representing two eyes and a mouth in a face), patients could correctly report the items on both sides of the screen. Without the "mouth" and "face-outline", patients ignored the item on the left. Further, work by Palermo and Rhodes (2002) suggests that upright and inverted faces may have different attentional resources. Subjects performed a sequential match-to-sample version of the part-whole task, but while viewing the upright test face were either asked to ignore two flanking faces or to say whether the flanking faces where pictures of the same person in different views. The flanking faces task interfered with holistic processing of the main task when the flanking faces were upright but not when they were inverted.

Practice and inversion effects. Both objects and faces show inversion effects on recognition memory. Objects also show large inversion effects in time (or accuracy) to name rotated objects (e.g., Jolicoeur, 1985; Lawson & Jolicoeur, 1998), but these effects disappear rapidly with practice (3-30 trials, McKone & Grenfell, 1999). Further, this is true even for objects which require discrimination of similar individuals rather than the more common basic level naming (approximately 95 trials, Tarr & Pinker, 1989). Training with inverted objects may even induce reversed inversion effects (i.e., better performance for inverted than upright). In a conference presentation, Husk, Sekuler, & Bennett (2004; Husk personal communication, August, 2004) report training subjects to discriminate upright or inverted houses using up to 4,000 trials in total. Before training there were no consistent inversion effects. After training there was an inversion effect for the subjects who had learnt houses upright and a reversed inversion effect for those that had learnt the houses inverted (about 13% in both cases). Note that this was still substantially smaller than the inversion effect for untrained faces (about 30% in this study).

There has been little research on training with inverted faces, particularly regarding whether training affects processing style. Haggbloom and Warnick (2003) showed, for example, that after training subjects can improve old/new recognition memory for inverted faces so that it is similar to the performance of another group on upright faces. However, this study did not test whether the processing of upright and inverted faces was the same. The only previously published study which tested processing style of inverted faces after training is the categorical perception study of McKone et al. (2001). They gave subjects up to 30,000 exposures to inverted faces but still failed to find holistic processing. It may be that the constancy or not of inversion effects over practice differentiates faces and objects (a full review on the subject of training with inverted faces is provided in Chapter 3).

1.2.6 Summary of behavioural studies.

The above sections have reviewed literature on behavioural findings for faces. Faces show disproportionate inversion effect in recognition memory compared to other objects. Further, as revealed by the composite, part-whole and spacing versus local change paradigms, upright faces seem to be processed in a holistic/configural manner including strong sensitivity to second-order relational information, whereas inverted faces are processed in a part-based or local feature manner. Not all the paradigms comparing upright and inverted faces have been used to compare faces and non-face objects. For the composite effect, the one object class tested to date (greebles) suggests no holistic processing for non-face objects. The part-whole effect is small but present for several object classes (i.e., the part-whole effect, like the inversion effect, is merely disproportionately large for faces rather than being present for faces and absent in objects). Spacing versus local changes have not been tested for objects in the same way that faces have, but note that this may be partly due to difficulty in defining what makes a spacing rather than a part change in, for example, a dog. It is also worth noting that in all the studies reviewed above, the objects, like faces, were compared at a within-class or individual level (i.e., comparing dog 1 vs. dog 2, in the same way that we would usually make judgments about Bill vs. Jim). This point will be returned to later.

Overall, behavioural findings show that upright faces are different from inverted faces and objects in terms of style of cognitive processing. The following sections will review differences (and similarities) in the way that objects and faces are processed in the brain. Sections will include brief reviews of studies looking at clinical patients, brain scans, hemispheric differences, and evoked response potentials (ERP) or magnetoencephalography (MEG).

1.2.7 Differences between faces and objects in neuropsychology.

Prosopagnosia is the disorder of being unable to recognise faces[2]. It was first reported by Quaglino & Borelli in 1867 (cited in Benton, 1990), and is often taken as evidence that faces are processed separately from other objects in the brain (e.g., Farah, 1996; McNeil & Warrington, 1993). However, it has been questioned whether prosopagnosia is truly a face-specific deficit, or whether the deficit applies generally to within-class discrimination (Faust, 1955, cited in Benton, 1990). Particularly in early studies, it was common to compare between class discrimination of objects (e.g., chair, vs. table) to within-class discrimination of faces (e.g., Bill vs. Jim), but it has been noted (e.g., Damasio, Damasio, & Van Hoesen, 1982) that this is not a fair comparison as individual faces are more similar to each other than objects such as a table and a chair are to each other. A single system for recognising both faces and objects, with partial brain damage distributed throughout this system, might show deficits for the harder (within-class) task of face discrimination, but not for the easier (between-class) task of object discrimination.

On the other hand, if the brain areas important for object processing are distinct from, but spatially close to, those for important for face processing, then a brain injury might often damage both areas together. Arguing against the within-class discrimination interpretation of prosopagnosia, apparently "pure" cases have been reported. For example, prosopagnosic WJ showed normal performance for identifying individual sheep faces, despite very poor performance for identifying individual humans faces (McNeil & Warrington, 1993; see also De Renzi, 1986). This result has recently been extended from patients with acquired prosopagnosia (e.g., as the result of a stroke or

[2] Note that prosopagnosics can sometimes recognise people by other features such as voice or hairstyle.

head-injury) to patients who were apparently born prosopagnosic (they have no history or sign of trauma). Duchaine, Dingle, Butterworth & Nakayama (2004) showed that the developmentally prosopagnosic subject Edward could be trained to recognise greebles at the same rate as normal controls, showing that his ability to discriminate similar individual objects was good. Moreover, a double dissociation between object and face recognition has been demonstrated. The "pure" object agnosic patient CK identified individual faces perfectly but was unable to recognise objects even at the basic level (Moscovitch et al., 1997; also see Humphreys & Rumiati, 1998). This rejects the idea that face processing is always more sensitive to brain damage than object processing.

The double dissociation between "pure" prosopagnosia and "pure' object agnosia would seem convincing evidence for a genuine separation between faces and objects. Despite this, some authors have quite recently argued for the within-class interpretation of prosopagnosia. Gauthier, Behrmann and Tarr (1999) varied level of categorisation for objects, including a between-class condition (e.g., comparing faces to other objects), a subordinate level condition (e.g., comparing a pigeon to an eagle) and an individual level condition (e.g., comparing eagle 1 to eagle 2). Noting that the tests of patients WJ and LH (suggested to be normal at making within-class discriminations of chairs and eyeglasses frames; Farah, Levinson, & Klein, 1995) were based only on accuracy measures with unlimited presentation duration, Gauthier et al. measured reaction time (RT) and also accuracy with limited presentation time. In simultaneous matching and match-to-sample tasks, two prosopagnosics (SM and CR) were affected by level of categorisation; age-matched controls were also affected, but the effect was disproportionately large for the prosopagnosics. Gauthier et al. interpret this finding as showing that prosopagnosics have a deficit at within-class processing for objects. Note, however, that the prosopagnosics were slower than controls even at between-class discrimination, making it difficult to compare the size of the categorisation level effects between subject groups as these were in relation to different baselines. In contrast to Gauthier et al.'s results, the developmental prosopagnosic Edward was within the normal range of RTs on two levels of a greeble verification task (identity and family), as well as a naming task (Duchaine et al., 2004). Second, using simultaneous matching and match-to-sample tasks de Gelder, Bachoud-Levi, and Degos (1998) found prosopagnosic patient AD showed a reversed inversion effect (i.e., inverted better than upright) for within-class discrimination of shoes. This pattern matched that found for faces in this patient (and patient LH, Farah, Wilson, Drain, & Tanaka, 1995). The reversed inversion effect for shoes also occurred for LH (de Gelder & Rouw, 2000), but was not found for a third prosopagnosic (RP in Rouw & de Gelder, 2002).

In summary, there is quite good evidence from the neuropsychology literature that faces and objects are processed separately in the brain. This is not without controversy, with some prosopagnosic patients also having deficits at making within-class discriminations with objects. However, note that drawing conclusions from this literature without reference to other areas of research may be misleading, as different patients have different levels and extent of damage (or in the case of developmental prosopagnosics, areas which have not been properly formed). This makes it hard to know which kinds of processing are linked from examining a single patient (see Duchaine, Nieminen-von Wendt, New, & Kulomaki, 2003, for a good discussion of this problem).

1.2.8 Face specific brain regions and neuroimaging.

As well as research on people unable to recognise faces (or objects) studies of brain processes have focussed on cortical activity in normal subjects. Using neuroimaging techniques, a "fusiform face area" (FFA; Kanwisher, McDermott, & Chun, 1997) has been located that is activated more by looking at faces than by looking at a range of other objects differing in basic category (activation is measured as an increase in Blood Oxygenation Level Dependent – BOLD – response)[3]. In arguing that this activation is genuinely face specific, rather than being attributable to within-class discrimination for faces, McCarthy, Puce, Gore, and Allison (1997) found strong FFA activation for different individual faces shown in a field of other objects, but not for individual flowers shown in a field of other objects. A possible criticism is that this difference could have been due to the passive viewing task leading subjects to spontaneously employ within-class processing for faces, but only basic level categorisation for flowers. However, using a one-back-matching task to ensure attention to individual identity for all stimuli, FFA activation has been confirmed to be stronger for faces than for hands (Kanwisher et al., 1997), and houses (Tong, Nakayama, Moscovitch, Weinrib, & Kanwisher, 2000). Similarly, higher activation for faces than houses, cars, flowers, or guitars has also been found on a within-class task where subjects have to make judgements about whether a stimulus is the same sub-type as a pre-specified target (e.g., pigeon rather than another bird; Grill-Spector, Knouf, & Kanwisher, 2004); and for individual Lepidoptera (butterflies and moths) on an old/new recognition task (Rhodes, Byatt, Michie, & Puce, 2004).

In contrast to this evidence, two studies by Gauthier and colleagues have argued that the FFA is a general within-class object processing area. In the first Gauthier, Anderson, Tarr, Skudlarski, and Gore (1997) found increased activation with subordinate level verification (e.g, Is this a penguin? for penguin vs. ostrich) compared to basic-level verification (e.g., Is this a bird? for penguin vs. goldfish). However, in this study the "FFA" was not individually localised; instead a large area of cortex was defined based on coordinates determined from previous studies. This means that the area being examined was not necessarily the same as the area that processes faces (McKone & Kanwisher, in press). In the second study, Gauthier, Tarr et al. (2000) individually localised the FFA, and replicated the result of increased activation with subordinate level verification. However, the amount of activation for within-class discrimination of objects is difficult to interpret given that no comparable task for faces was included; the only comparison made was between the difference between faces and objects (localiser task) and the difference between basic-level and subordinate-level discriminations.

Overall, there seems to be good evidence that the FFA is more activated by faces than other objects. Even when processing is at an individual level the activation for faces is still stronger than that for objects (Rhodes, Byatt, et al., 2004). The only object class for which there seems to be much more activation than baseline is birds (Grill-Spector et al., 2004). As Grill-Spector et al. themselves point out, this may be because birds, like faces, have heads.

A second question about the FFA is then in what way is it involved in face processing? The fact that the FFA is quite strongly activated by cat faces and by human eyes without the rest of the face (Tong et al., 2000) may suggest that any "faceness" is enough to activate the area to some extent. This suggestion is further supported by the fact the area is still strongly activated by inverted faces, although less than for upright faces (Kanwisher, Tong, & Nakayama, 1998). Importantly, however, recent evidence from Grill-Spector et al. (2004) in which they correlated activation with behavioural

---

[3] Other areas such as an Ocipital Face area (OFA) are sometimes reported, but as the focus in the literature has been on the FFA, focus in this thesis will also be on the FFA.

responses on a trial-by-trial basis, showed that the FFA is involved in individual face identification as well as just face detection.

The question of what activation actually means in these studies is also an important one. Generally results are discussed as though more activation in an area indicates better tuning for processing a particular stimulus class. However, less activation can be found for familiar than unfamiliar faces in the FFA (e.g., Rossion, Schiltz, & Crommelinck, 2003), and for high-frequency than low-frequency words in word areas (e.g., Kronbichler et al., 2004), suggesting that the predicted response may not always be clear. Although not a problem in most of these studies, it is also worth noting that attention modulates the FFA response (Wojciulik, Kanwisher, & Driver, 1998).

1.2.9 Hemispheric differences in processing style.

For faces, standard behavioural findings are that (a) upright faces are recognised better with initial presentation to the right hemisphere (left visual field) than with initial presentation to the left hemisphere (e.g., Rhodes, 1993; Rhodes et al., 1993; Watanabe, Kakigi, & Puce, 2003), and (b) inversion effects are larger with right hemisphere presentation than with left hemisphere presentation (McKone, 2004; Rhodes, 1993; see also Le Grand, Mondloch, Maurer, & Brent, 2003). Further, Parkin and Williamson (1986) showed that subjects were faster to make a face/non-face decision (Mooney faces vs. random blobs) when initial presentation was to the left visual field/right hemisphere, but were faster to say that features were incorrect (e.g., nose replaced with a picture of a telephone) when initial presentation was to the right visual field/left hemisphere. They interpret this result as showing that the right hemisphere is better at holistic processing, but that the left hemisphere is better at part-based processing. In terms of imaging, Rossion, Dricot et al., (2000) supported this conclusion using a part-whole task. They found that the right FFA showed stronger activation when subjects made a judgement about whole faces than when they made a judgement about isolated face parts, while the homologous areas in the left hemisphere (left FFA) showed the reverse effect. This pattern was not replicated for houses; instead, there were no significant differences in activation between judgements about whole houses and parts of houses in either hemisphere.

1.2.10 Temporal differences in brain response to faces and objects.

Studies of brain activation using neuroimaging, although allowing examination of spatial differences, fail to capture possible temporal differences in neural processing of faces and objects. For this, event-related potentials (ERP) or magnetoencephalography (MEG) are needed. Using ERPs, it has been reported that there is a negative component at approximately 170 milliseconds which is face-specific over some electrodes (T5 and T6 which are usually located over the middle temporal gyrus in the left and right hemispheres respectively). The N170 is larger in response to faces than to individual items from a range of stimulus classes including cars, furniture, human hands, animal faces (Bentin, Allison, Puce, Perez, & McCarthy, 1996), shoes, houses and greebles (Rossion, Gauthier et al., 2000). Bentin et al. (1996) suggested that the N170 might best be described as sensitive to eyes rather than whole faces, as it is also very strong for eyes presented alone, inverted faces and faces with the position of the features scrambled. The effects are actually slightly different with the amplitude for inverted faces being enhanced and the effect slightly delayed (approximately 10 ms, replicated by Rossion, Gauthier et al., 2000). Eyes alone also produce increased amplitude (Bentin et al., 1996).

Using MEG, Liu, Harris and Kanwisher (2002) found an M170 (M as in MEG) that was larger for faces than for houses on an identification task. They further found that this M170 was larger to a face which had had the features blacked out than to a face where the features had been scrambled, suggesting that configuration was more important than local feature shape (including eyes). This was contrasted with an M100 which showed a preference for the scrambled real face (correct features) over the face with features blacked-out (correct configuration).

Thus, there is some kind of processing at 170 ms that differs for faces and objects, and even for inverted faces. It seems likely that this is involved in face identification, whereas an earlier process is involved in face detection.

1.2.11 Summary of brain studies

Studies of face-processing in the brain support the findings from the behavioural literature that there are differences between face and object processing. Studies with "pure" prosopagnosics and object agnosics show that objects and faces are processed by different areas of the brain. This is supported by neuroimaging studies which show that the FFA is activated more by faces than other objects, even when within-class discrimination is used. Studies with MEG and ERP also show that the time-course for processing of faces and objects differs.

1.3 Summary of the evidence for 'special' processing of (upright) faces

Overall, the review in Chapter 1 has presented substantial evidence that faces are processed differently from other objects. Behaviourally, there are both quantitative differences between faces and objects (inversion effects and part-whole effects) and qualitative evidence that only faces are holistically processed (composite effect, peripheral identification). There is also evidence that faces and objects are processed largely in different brain areas, and with different time courses. These distinctions are not uncontroversial, however, with slight effects such as the part-whole effect for objects, or some activation in the same brain areas taken by some researchers as evidence that faces and objects are not processed differently when within-class discrimination is required.

I note that the only real exception to the special status of faces might be human bodies. Reed, Strong, Bozova, and Tanaka (2003) found the same size inversion effect for faces and bodies for both accuracy and reaction time[4]. There is also a part-whole effect for bodies. Seitz (2002) found an advantage for whole bodies over parts (e.g., arms) that was the same size as that for faces when the "whole" condition for faces and bodies was matched. There was also a "whole" advantage for bodies in 8 and 10 year old children, which appears to be as large as or larger than the effect for faces (for children the whole condition for faces and bodies was not matched). Although the FFA is not activated by hands, there is an area of the brain which is more activated by body parts than other objects including whole faces (Downing, Jiang, Shuman, & Kanwisher, 2001). Thus, it may be that similar cognitive processing of bodies to faces is occurring, but in a different cortical area. This is an idea which needs more investigation, but not one which is covered by this thesis.

[4] Note that the accuracy effect was less than half the size of that usually found for faces, 7% in this study compared to an effect for faces of usually 15-25%. This may be partly due to ceiling effects and partly due to the fact that distractor stimuli "differed on one or two features, such as facial hair" (p.305), lessening the need for holistic processing.

The next chapter discusses theories of why the results reviewed above might occur. That is, how the pattern of similarities and differences in object and face processing might have arisen. Also discussed are possible origins of "special" processing for faces, and theories of how individual faces might be represented.

CHAPTER 2: THEORIES OF FACE RECOGNITION

This chapter introduces theories of face recognition considered from three perspectives. In Section 2.1, I first describe theories of why upright face processing appears to be different from inverted face processing and object processing; these theories include the within-class discrimination hypothesis, the expertise hypothesis and domain specificity. In Section 2.2, I then review the possible origins of special face processing consistent with the domain-specific view. Finally, in Section 2.3, I describe theories of how faces are represented as individuals, and coded with respect to an average face, with particular emphasis on the concept of face-space.

2.1. Theories of why (upright) faces appear special

In previous literature, three primary theories of why faces appear to be special have been proposed. These are referred to here as the within-class discrimination hypothesis, the expertise hypothesis, and the domain specificity view. I now briefly describe each theory and summarise how and where this thesis addresses them (detailed reviews of the relevant empirical literature are given in other Chapters).

2.1.1 Within-class discrimination is sufficient for holistic processing.

One potentially important difference between faces and objects is the preferred level of identification. Humans naturally process faces at the level of individual identity (e.g., Sam vs. Bob); that is, in a within-class fashion. In contrast, objects are usually identified at merely the basic level (e.g., Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), such as chair versus dog, or at a subordinate but not within-class identity level, such as poodle versus labrador (but not poodle 1 vs. poodle 2). An early idea was that this difference in processing level might be the source of apparently special processing for faces (e.g., Faust, 1955, cited in Benton, 1990; Damasio, Damasio, & Van Hoesen, 1982).

This theory would then predict that any object class should be processed in the same way as faces if the task requires identification at a subordinate level. Although some quite recent articles still seem to argue in favour of this hypothesis (e.g., de Gelder & Rouw, 200; Gauthier, Tarr et al., 2000; Tarr, 2003), it is generally agreed that it has been disproved. All the evidence reviewed in Chapter 1 supports the view that the within-class discrimination hypothesis is not viable. In terms of neuropsychology (Section 1.2.7) there are some prosopagnosics who have deficits at making within-class discriminations for both faces and objects, but there are also some patients who have "pure" face deficits or object deficits. Similarly in the area of neuroimaging (Section 1.2.8), although within-class categorisation sometimes increases activation for non-face objects in the FFA (Rhodes, Byatt, Michie, & Puce, 2004), in general this is not the case (Grill-Spector, Knouf, & Kanwisher, 2004) and the BOLD response in the FFA remains much less than for faces in all studies. Finally, the behavioural studies reviewed all used tasks which required within-class discrimination of objects (e.g., recognition memory for a series of houses). Despite this, inversion effects were smaller for objects than for faces (Section 1.2.1), the part-whole effect was smaller for objects than for faces, and the composite effect was absent for greebles (Section 1.2.4). All these findings argue that the special holistic processing which occurs for faces does not occur for individual discrimination of objects.

2.1.2 Expertise hypothesis: within-class discrimination of objects-of-expertise elicits holistic processing.

In addition to preferentially categorising faces at an individual level, humans are also much more experienced at making these fine discriminations for faces than for other objects. That is, we are expert at making individual face discriminations, but are usually not expert at making individual object discriminations. The expertise hypothesis (Diamond & Carey, 1986; Gauthier & Tarr, 1997; Meadows, 1974) argues that this difference can explain the empirical differences found between faces and objects. It then predicts that face-like processing for objects should occur when three conditions are met (Diamond & Carey, 1986). First, like faces, individual exemplars of the object class should share a first-order configuration and differ on second-order relationships (i.e., have the same basic parts and differ in spacing between those parts). Second, as for faces, the task should require that objects are recognised at an individual level. Third, as for faces, subjects should be expert enough with the stimuli to make use of subtle second-order differences between individuals.

Studies testing the expertise hypothesis have used experts of two very different levels of experience. Experiment trained experts are given approximately ten hours of training with individual members of an object class (these studies have almost exclusively used greebles as stimuli). An obvious limitation of studies with experiment trained experts is that, while demonstrating the emergence of a face-like effect in this time-frame would be very interesting, a lack of effect does not mean that it could not be found with greater experience. Addressing this issue, real-world experts with many years of experience have also been tested. These real-world experts have included groups such as dog-show judges, car enthusiasts and bird-watchers.

The literature review of empirical studies testing the expertise hypothesis is provided in Chapter 5, and to some extent Chapter 3. As will be shown in these chapters, there is currently very active debate about the expertise hypothesis. While many authors seem persuaded that "special" aspects of face recognition are based on expertise (e.g., Elgar & Campbell, 2001; Le Grand, Mondloch, Maurer, & Brent, 2004; Reed, Stone, Bozova, & Tanaka, 2003), others are less convinced (e.g., Duchaine, Dingle, Butterworth, & Nakayama, 2004; McKone & Kanwisher, in press; Tanaka, Kiefer & Bukach, 2004). My opinion of this literature, especially arising from the most recent studies (e.g., Grill-Spector et al., 2004; Rhodes, Byatt, et al., 2004), is that convincing evidence in support of the expertise hypothesis has not been demonstrated. Indeed, I believe the published evidence is generally against it, and my own data (Chapters 3 and 5) will also fail to show any development of face-like processing with expertise.

2.1.3 Domain specificity: Faces *per se* are special.

Some authors (e.g., Kanwisher, 2000; McKone & Kanwisher, in press; Rhodes, Byatt, et al., 2004) have argued that faces *per se* are special. According to this domain specific view, faces are the only stimulus class to strongly activate the FFA, and may be the only stimulus class that can be holistically processed (a possible exception being bodies). This theory predicts that neither within-class discrimination nor expertise is sufficient to produce face-like processing of other objects. Note that, logically, the domain specificity view can only be supported by eliminating the other theoretical alternatives.

Also note that the domain-specificity idea itself says nothing about the origins of special processing for faces (beyond saying that it does not arise from generic expertise). Given my suggestion that current evidence for the expertise hypothesis is weak, I consider it valuable to examine the possible origins of holistic processing within the domain specific view, to assess whether it is a viable alternative. This is done in Section 2.2.

2.1.4 How the present thesis addresses theories of why faces appear special

In summary, three general theories of why faces might be special have been suggested. The two considered most viable in the current literature are the domain specificity hypothesis (e.g., Grill-Spector et al., 2004; Kanwisher, 2000; McKone & Kanwisher, in press; Rhodes, Byatt, et al., 2004) and the expertise hypothesis (e.g., Gauthier, Curran, Curby, & Collins, 2003; Gauthier & Nelson, 2001; Gauthier & Tarr, 1997, 2002). Domain specificity is the idea that faces are somehow special per se. The contrasting expertise hypothesis suggests that faces are simply the class with which we have most experience at making within-class discriminations. That is, extensive experience is suggested to be sufficient for face-like processing to occur for any non-face objects. The amount of experience required to produce face-like processing has been suggested to be approximately 8-10 hours of laboratory training in some studies, or many years of real world experience in others. In the present thesis, both the "8-10 hours" and "many years" timescales for expertise are directly tested as part of the first aim of the thesis (to test whether face-like processing is really domain specific). In particular, Chapter 3 tests experimental training with inverted faces, while Chapter 5 tests real-world expertise with labrador dogs using dog-experts of many years' experience.

The third theory is that within-class processing alone, even without expertise, is sufficient to produce face-like processing for objects. This theory has little empirical support, but has still been taken seriously quite recently (e.g., Tarr, 2003). Results presented in Chapter 4 (comparing Labrador and face processing in novices) add to the literature relevant to this hypothesis.

2.2. Possible origins of special face processing.

Within the domain specificity view, many authors have suggested an origin of special processing for faces in terms of an innate representation of basic face structure. They further suggest that such a representation might have arisen from a system for recognising/imprinting on members of one's own species (conspecifics) that is maintained across different species. For example, Morton and Johnson (1991) have suggested an innate representation of the basic structure of a face (three blobs in the positions of eyes and nose/mouth), which they propose might be coded at a subcortical level (see also Mondloch et al., 1999). Another possibility is that there is no innate representation, but faces are the first thing that one regularly sees (or the first thing that is regularly placed at a distance at which infants can focus). Either of these alternatives allows that there may be a critical period for the development of normal face recognition. That is, there may be a period in development in which it is important to have specific face input, in the same way that kittens raised without any early exposure to vertical (or horizontal) lines are insensitive to horizontal (or vertical) edges later in life (Blakemore & Cooper, 1970; Hirsch & Spinelli, 1970).

Two types of empirical studies are relevant in assessing these possible origins of special face processing in adult humans. The first are studies including both normal and visually deprived infants and children. The second are studies on whether face recognition has special developmental status in other animals, especially those closely related to humans in evolutionary terms. Both will be briefly reviewed.

2.2.1 Developmental studies in humans.

Face preference in normal infancy. If there is an innate face representation, infants might prefer face-like stimuli compared to those that are less face-like. Morton and Johnson (1991) suggest that there is an innate mechanism that causes infants to track face-like stimuli in early infancy. This would ensure that infants are exposed to faces during a (postulated) critical period, so that they develop normal face processing (see also de Haan, Humphreys, & Johnson, 2002). There is empirical evidence to support this idea. Newborns less than one hour old tracked a schematic face stimulus further than a scrambled schematic face stimulus (Johnson, Dziurawiec, Ellis, &

Morton, 1991). Newborns less than two hours old preferred three blobs in the correct placement for eyes and nose to the same blob configuration upside-down (Johnson et al., 1991; Mondloch et al., 1999). This second preference seems to disappear by six weeks of age (Mondloch, et al., 1999) with the preference for schematic faces over scrambled schematic faces also disappearing by three months of age (Johnson et al., 1991). Newborns less than 4 hours old also looked longer at their mother's face than a stranger's face (in this experiment, real heads rather than photographs were used), but this preference disappeared when the hair and face-outline were covered with a scarf tied under the chin (Pascalis, de Schonen, Morton, Deruelle, & Fabregrenet, 1995).

Some authors suggest that infants simply prefer aspects of stimuli which happen to be more frequent in face-like stimuli than other stimuli tested (see Simion, Cassia, Turati, & Valenza, 2001, for a review). However, this fails to explain why infants should prefer stimuli that happen to be more face-like. Unless an explanation can be given, the parsimonious suggestion would seem to be that infants prefer faces. Cassia, Turati, and Simion (2004) argue that it is information in the top-half of the visual field (i.e., more elements), rather than faceness per se which infants prefer. Cassia, et al. compared looking preference in infants less than 4-days-old for pairs of stimuli in several combinations. When an upright face was contrasted with the same head with the internal features inverted as a group, infants preferred the upright feature group to the inverted feature group. When a scrambled face with more visual information in the top-half (two eyes turned sideways and off-set to the right, above the nose turned sideways, above the mouth off-set to the left) was contrasted with a scrambled face with more information in the bottom-half (the mouth off-set to the left, above the nose turned sideways, above two eyes turned sideways and off-set to the right) infants preferred the scrambled face with more visual information in the top-half. When the scrambled face with more visual information in the top-half was contrasted to the normal upright face infants had no preference. However, there are some problems with the conclusion from this study that infants simply prefer information in the top-half of the visual field. The most obvious is that, at low spatial frequencies, the face stimuli which had more information in the top-half also look more face-like than those which have more information in the bottom-half. Another criticism is that although these stimuli have more information in the top-half, normal faces seen as part of whole heads do not (i.e., the eyes are usually situated approximately half-way up the head, meaning that there are more elements in the lower half). However, the problem of stimuli cropped such that more information is present in the top-half is also true of the original Johnson et al. (1991) schematic faces.

To establish that there is an innate preference for faces several things would be needed. A very positive aspect of the Cassia et al. (2004) study is that they based their stimuli on pictures of real faces, thus replicating a preference for upright feature groups over inverted feature groups with more naturalistic stimuli. As well as using real face stimuli, another important test of innate representation would be to test that infants are more interested in faces than other objects. It seems most likely that they would prefer the item for which there is an innate representation; however, it is also possible that the other item would be seen as more novel and hence attract more interest. An even stronger test would therefore be to compare infants' ability to discriminate individual faces to their ability to discriminate individual non-face objects: an innate capacity for faces would predict better discrimination for faces than objects. To my knowledge neither of these have been done with infants less than several months old.

Studies in infants a few months to a year old have examined whether they can discriminate non-face objects. Quinn and Eimas (1996) familiarised three to four month old infants with pictures of cats, and then showed pairs of a novel cat and a novel dog either as whole animals, as bodies without heads, or as heads without bodies. Dogs (the unfamiliar stimulus class) were preferred in the whole body and head-only conditions, but there was no preference in the body-only condition. This suggests that dogs could be distinguished from cats on the basis of head shape, but not body shape. This only occurred for the upright orientation; in the inverted orientation, dogs were preferred in all conditions. Results from Bonatti, Frot, Zangl and Mehler (2002) suggest that 10-month-old infants can distinguish human-like objects (dolls) from other objects (including toy dogs faces), and distinguish dolls from one another (although this result was less strong) but not distinguish non-face-like objects (e.g., black motor car vs. red wax strawberry) from one another. In this experiment infants were shown two objects (or one object and one doll) one at a time, and then subsequently shown either both the objects together or only one of the objects. If infants looked longer in the single object condition, it was concluded that they were surprised that there was one not two objects, and hence that they had been able to tell the two objects apart. The finding that 10-month-old infants can discriminate between stimuli containing faces, but cannot perform even between-class discriminations for stimuli which do not contain faces is hard to reconcile with the idea that faces are not special, given that infants can discriminate two similar looking unknown males (a difficult within-class discrimination) presented in 3/4 view, at 5 months old (Fagan, 1979).

A preference for faces in infancy may of course be tuned by experience. Using a habituation paradigm, Pascalis, de Haan and Nelson (2002) showed that six month old infants can distinguish between individual monkey faces, as well as between individual human faces (they look longer at the novel face for both species). This was not true for 9-month-old infants and adults, who could distinguish only human faces.

Overall, these studies suggest that before six months infants may be interested in faces generally, and are better able to discriminate faces than to discriminate other stimuli, but also that the system is tuned by experience to differentiate human faces. The results are consistent with the idea that infants have some innate representation of faces, or at least develop such as representation very early.

Holistic processing in normal development. The question of whether children holistically process faces, or at what age they do so, is important to the question of domain specificity for two reasons. First, it was in this context that the idea that holistic processing may require extended experience to develop (the expertise hypothesis) was first suggested. The second related reason is that tests of holistic processing in children can help to put an upper-limit on the amount of training needed to develop holistic processing if it is simply a matter of generic expertise.

Early studies by Carey and Diamond suggested that holistic processing was not mature until approximately ten years of age. This was on the basis that children six- or eight-years-old did not show adult like inversion effects (Carey & Diamond, 1977) and were more likely to judge faces as the same based on paraphernalia (e.g., hats, glasses) than on identity (Diamond & Carey, 1977). However, these studies were refuted by later evidence.

A number of more recent studies argue that holistic processing is functional at much earlier ages. Carey & Diamond (1994) showed that there was a composite effect (Young, Hellawell, & Hay, 1987) indicative of holistic processing at ages 6, 10 and adult. Similarly, Tanaka, Kay, Grinnell, Stansfield, & Szechter (1998) showed that parts were recognised better in the context of whole faces than when isolated (a part-whole effect; Davidoff & Donnelly, 1990; Tanaka & Farah, 1993) for children 6-, 8- and 10-years-old. This has also been extended to children only 4-years-old (Pellicano & Rhodes, 2003). Further, Pellicano, Rhodes and Peters (submitted) included Tanaka and Sengco's (1997) version of the task in which the part is presented in a version of the original face with a spacing change made to it (the transformed version) as well as in the context of the original face. Four-year-olds were better at the task when the original face was presented compared to the transformed face, suggesting that they did code spacing information. Findings of holistic processing at these ages also comes from work by Gilchrist and McKone (2003) for 6- and 7-year-olds, and McKone and Boyer (submitted) for 4-year-olds. These studies used differences in spacing or features and asked questions about distinctiveness in a paradigm based on Leder and Bruce (1998). Importantly, Gilchrist and McKone showed evidence of holistic processing in 6- and 7-year-olds (better ability to remember/discriminate spacing changes in upright faces than inverted faces) when baseline performance was matched to that of adults. This suggests that ability to code spacing changes was quantitatively as well as qualitatively mature at this age (although see Mondloch, Geldart, Maurer, & Le Grand, 2003; Mondloch, Le Grand, & Maurer, 2002).

The above paradigms have not been tested with children younger than 4-years-old (because of difficulties in explaining the task and so on). However, Cohen and Cashon (2001) have at least shown that 7-month-old infants can integrate information from the external (e.g., hair) and internal aspects of the face. In a habituation paradigm, infants looked longer at a face containing internal features from one previously seen face and external features from another previously seen face, than at a truly "old" face (this is also consistent with the finding of Pascalis et al., 1995). Geldart, Maurer and Henderson (1999) have also shown that 5-month-old infants can differentiate faces with the features moved (as a group) to a higher than average position from faces with features moved to a lower than average position (infants looked longer at faces with features higher than average), suggesting that the position of the features can be encoded to some extent at 5 months old (see also Thompson, Madrid, Westbrook, & Johnston, 2001).

Overall, evidence clearly supports holistic processing in children as young as 4-years-old on standard face tests. This is an important finding for tests of the expertise hypothesis as it sets an upper limit of 4 years of experience for developing holistic processing, if expertise alone is responsible for the usual finding of differences in face and object recognition.

<u>Effects of early visual deprivation and brain damage.</u> Logically, whether four years' experience (or any other amount) is sufficient to develop holistic processing for objects other than faces also depends on whether there is a critical period in infancy for

developing holistic processing for faces. This question has been addressed by studies of patients with early visual deprivation, or early brain damage.

Studies by Le Grand and colleagues investigated the effect of early visual deprivation by testing the face recognition abilities of patients born with dense cataracts which were later removed. These patients had at least eight years of normal vision after the removal of the cataracts, but no detailed pattern vision to one or both eyes for the first 3- to 6-months of life. In a number of studies these patients were shown to have normal ability to match faces on direction of eye-gaze, lip-reading and expression (Geldart, Mondloch, Maurer, de Schonen, & Brent, 2002) as well changes to features (Le Grand, Mondloch, Maurer, & Brent, 2001). They were, however, severely impaired at detecting changes in spacing (Le Grand et al., 2001) and matching identity over changes in head position (Geldart et al., 2002). Further, they did not show Young et al.'s (1987) composite effect (the difference between aligned and unaligned composite faces), and this was not due to worse overall performance than controls (Le Grand et al., 2004). The hemisphere which is deprived is also important. Le Grand, Mondloch, Maurer, and Brent (2003) showed that only patients with bilateral or left-eye cataracts had deficits on the spacing task. This equates in infants to right hemisphere deprivation, that is, the hemisphere shown to be more important in face processing in adults (Section 1.2.9). These studies strongly suggest that there is a critical period for developing normal holistic face processing. This idea is consistent with studies of developmental prosopagnosics and of patients who have the area of the brain usually used for face processing damaged at a very early age (e.g., one day old, Farah, Rabinowitz, Quinn, & Liu, 2000).

One theoretical possibility is that any stimulus class might be holistically processed if it is seen enough during the critical period, but that infants' visual environment is set up in such a way that this usually only occurs for faces (most parents would prevent the family dog from spending much time with its face close to their baby's face). Thus, the strongest evidence of a face-specific innate mechanism would be if newborns could distinguish individual faces (which there is some evidence that they can) but not distinguish individuals of another class (e.g., dogs).

2.2.2 Face-recognition in non-human animals

The question of whether faces are special per se, and whether this has an innate basis, may also be considered from the point of view of non-human animals. Morton & Johnson (1991) have argued in general for an innate mechanism which allows any species to more easily recognise its own kind (i.e., conspecifics). A study by Pascalis and Bachevalier (1998) showed that in a habituation paradigm rhesus monkeys spent more time looking at a novel face than a previously seen face of their own species, but did not discriminate between a novel and previously seen human face, even though these monkeys had been raised by humans. Humans in this study looked longer at a novel human face and did not differentiate between novel and previously seen monkey faces. With 8- to 9-year-old chimpanzees, also raised by humans, Parr, Dove and Hopkins (1998) found an inversion effect on a sequential matching task that was larger for chimpanzee faces than human faces, but absent for capuchin monkey faces or cars.

Monkeys also have a face-specific area in the brain in the same way as humans do. Tsao, Freiwald, Knutsen, Mandeville, and Tootell (2003) showed that there is an area in the macaque brain which appears to be homologous to the FFA. It is in a similar location in the macaque brain to that of the FFA in the human brain. The area is more active to faces (human and macaque) than to other objects or to Fourier phase scrambled textures (made from faces or objects). The area was also more active to macaque faces than to human faces. (The FFA of the human subjects was equally activated by the two kinds of faces.) Single-cell recordings in nonhuman primates also suggest that there are cells in the Superior Temporal Sulcus which are more strongly activated by faces and bodies than by other objects (e.g., Wachsmuth, Oram, & Perrett, 1994) and that these are tuned to specific orientations (Ashbridge, Perrett, Oram, & Jellema, 2000).

Conspecific recognition also shows some similarities between humans and other non-human animals. Sheep are better at recognising sheep faces of their own breed than another breed, and better at recognising upright than inverted sheep faces (Kendrick, Atkins, Hinton, Heavens, & Kevern, 1996). Further, sheep discriminate upright sheep faces predominately using the right hemisphere (Broad, Mimmack, & Kendrick, 2000). Right hemisphere use in conspecific recognition also seems to occur for 3-day-old chicks (Deng & Rogers, 2002). These results are consistent with the idea that animals at many levels of the evolutionary tree may have special ways of recognising conspecifics.

Note that none of the studies mentioned in this section directly test "special" face processing in the sense of holistic processing. To my knowledge, no one has conducted direct tests of holistic processing of faces (e.g., the part-whole or composite tests) in non-human animals. This would obviously be important in understanding whether the face recognition abilities shown by non-human animals are truly similar to those shown by humans.

2.2.3 The origins of special face processing: Summary.

Studies on infants and non-human primates reviewed above suggest that faces of one's own species are given some special status. Whether this has an innate basis for faces *per se*, or is due to the fact that faces are the only objects seen extensively during an early critical period, is not known. It is also important to note that evidence of holistic processing for upright, but not inverted faces, in adults could be explained by either view. That is, within a domain-specificity theory, the large inversion effect in adults could be due to either an innate mechanism coding the basic structure of an upright face, and/or due to the fact that not enough inverted faces were seen during a critical period.

2.3. Individual representations of faces and face-space

This section discusses aspects of how individual faces might be coded within the system that represents faces, and how the idea of a "face-space" might be integrated with the largely independent literature on holistic processing for faces. By the "system that represents faces" I mean the stage of processing that deals with perceptual representations rather than, for example processing people's names, semantic information (e.g., a person's job), or expression. In the Bruce and Young (1986) model of face recognition (see Bruce, Burton, & Craw, 1992, for a later version), this would be the stage dealing with identifying the face based on appearance. In this model, other aspects of face recognition such as expression are separable from the stage which deals with facial appearance/identity.

2.3.1 Representing individual faces in "face-space"

A common idea is that appearance of individual faces is coded with respect to an average, norm, or prototype. Theoretically, this idea is usually instantiated in the idea of "face-space", as proposed by Valentine (1991; also see Valentine & Bruce, 1986). This is a theoretical multidimensional space in which individual faces are coded on physical dimensions that usefully describe real-world variability between faces (e.g., face width, length of nose, etc.). A face lying at the centre of face-space is exactly average on all dimensions. Faces that are more atypical on one or more dimensions lie further from the centre on those dimensions. An example with only two dimensions is shown in Figure 2.1.

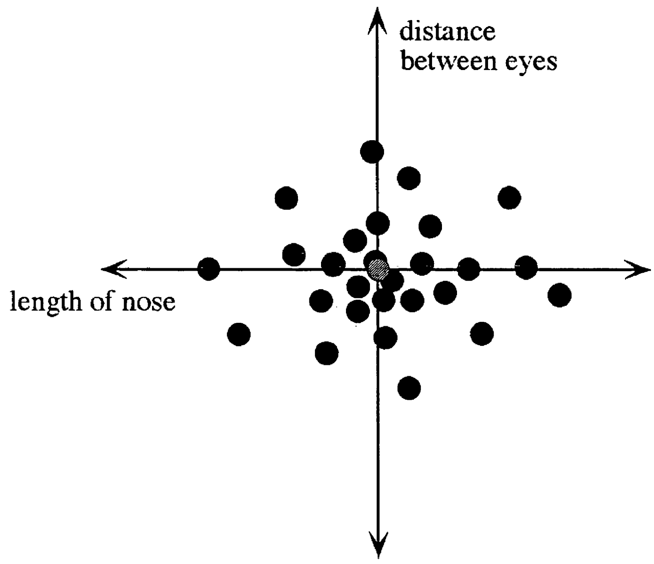


Figure 2.1. The heuristic of face-space, showing of exemplar faces (black dots) distributed on two possible dimensions. The average face is shown at the centre (grey dot).

The dimensions of face-space are usually thought of as not being defined *a priori*, but instead are presumed to be determined from the range of faces to which one has been exposed. Empirical evidence shows that prototypes of faces can be formed quickly and easily. Bruce, Doyle, Dench and Burton (1991) used a manipulation in which the internal features of a face were moved as a block with respect to the face outline. In a memory test, subjects were asked to choose which of a pair of faces they had previously rated for sex and approximate age. When the pair consisted of the unseen prototype face (or central feature version) and another unseen exemplar, subjects picked the prototype as seen on 82% of trials (Experiment 4). When the pair consisted of the unseen prototype and a previously seen exemplar, subjects chose each equally often (Experiment 5). This was after rating only four versions (exemplars) of each face. Infants less than 78 hours old are also able to quickly form prototypes from experienced faces. Walton and Bower (1993) showed that after seeing four faces (for less than a minute in total) infants preferred an average of these four previously seen individuals to an average of four previously unseen individuals.

Valentine (1991) deliberately left the exact nature of the underlying dimensions of face-space open. It is still unknown what the dimensions are, but two approaches have been used to determine what they might be. The first is to use a multidimensional scaling approach such as that used by Rhodes (1988). In this study, face pictures were rated on 31 scales ranging from things such as hairstyle and presence of freckles to forehead shape (e.g., receding) and eye-setting (e.g., protruding). Measurements of the face pictures were also taken and included things such as eye width, nose length, and lip thickness, as well as distances between features (39 measurements in total). The faces were then paired for similarity by a second group of subjects who paired the two most similar faces, then the next two most similar, and so on. (This method required far fewer possible pairings than would be required for rating similarity of all possible pairs.) Multi-dimensional scaling (MDS) was used on these similarity pairings and showed that three dimensions gave a good fit to the data. The ratings scales and measurements were then regressed against these three dimensions so that the dimensions could be interpreted. One dimension seemed to be linked to weight, with another linked to weight, hair-length and several measurements involving mouth-shape. The final dimension was linked to many measurements and ratings involving eye-shape, as well as nose-shape and a rating of the age of the face. Interestingly, eye-height in the face (measured to the chin), nose-length, and the relative positions of the eyes, nose and mouth in the face did not load on any of the three general dimensions.

An alternative approach to determining the kind of dimensions which might be coded in face-space is to use principle components analysis (PCA). In this kind of analysis the pixel-by-pixel intensity of a set of faces is scanned. From these values, eigen-vectors are calculated on the basis of which aspects of the face explain the most variance in the face pictures. Eigen-vectors themselves do not look like anything, but they can be added to or subtracted from an average face so that their effect may be seen. Early eigen-vectors (those that explain large amounts of variance) seem to code very general factors like masculinity/femininity. Individual identity is coded by eigen-vectors which each explain much less of the variance (e.g., O'Toole, Abdi, Deffenbacher, & Valentin, 1995). These two ways of exploring which aspects of faces might be coded in face-space differ in that PCA works from a set of physical faces used as input, whereas MDS is based subjects' perception of faces.

2.3.2 Empirical results commonly explained with reference to face-space.

The heuristic of face-space has been used to explain a number of important empirical findings. The distinctiveness effect refers to the fact that faces rated as "standing out more in a crowd" (distinctive faces) are recognised faster, but are classified more slowly as faces, compared to typical faces (e.g., Valentine & Bruce, 1986). This is usually explained in terms of face-space as follows (although see Burton & Vokey, 1998). Faces which are more distinctive lie further away from the norm (or centre) on one or more dimensions. The more typical faces cluster around the centre. Faces which are more distinctive therefore have fewer near-neighbours in face-space, making them less confusable and easier to recognise in a memory task. Typical faces are, however, classified faster as faces than are distinctive faces, because the typical faces are closer to the norm (Valentine, 1991).

Another similar effect is that of averageness on attractiveness. Faces created by mathematically averaging many other faces are rated as more attractive than the component faces (e.g., Rhodes, Sumich, & Byatt, 1999). This has also be explained in terms of face-space, by suggesting that faces closer to the centre of face-space are seen as more attractive.

Caricature effects refer to that fact that faces can be made more recognisable by distorting them further away from the average along a trajectory drawn from the centre of face-space to the individual. This has the most effect on features which are not average to begin with; for example, a face with a larger than average nose would have its nose made even larger. It is also possible to make "anticaricatures", that is, faces which are closer to the average than the original along the same trajectory. Empirical results show that making caricatures of faces causes them to be recognised faster than a veridical line drawing, whereas making anticaricatures of faces causes them to be recognised slower than veridical line drawings (Rhodes, Brennan, & Carey, 1987). Caricature effects can be explained in terms of face-space in a similar way to distinctiveness effects; that is, moving a face further from the average means that there are fewer faces to confuse it with. Rhodes et al. (1987) suggest that this advantage offsets the negative effect of the caricature not quite matching the veridical face.

The cross-race deficit refers to the finding that subjects are often better at recognising individuals of their own race than individuals of another race (e.g., Rhodes, Tan, Brake, & Taylor, 1989). There may even be more activation in the FFA for own than other race faces (Golby, Gabrieli, Chiao, & Eberhardt, 2001). The memory effect is explained in terms of face-space by suggesting that other race faces are coded using dimension values that are salient to that race but not necessarily to that individual (Valentine, 1991). Thus, other race faces will form a cluster some distance from the centre (i.e., some distance from the norm in a norm-based model). Other race-faces are then difficult to tell from one another although it is easy to tell that they are from another race.

Finally, adaptation aftereffects for faces may also be relevant to discussions of face-space. These will be discussed more fully in Chapter 6, but, briefly, adapting to a male face can make a sex-neutral face look female (Webster, Kaping, Mizokami, & Duhamel, 2004), and adapting to a radially expanded face makes subjects choose a slightly contracted face as both most normal and most attractive (Rhodes, Jeffery, Watson, Clifford, & Nakayama, 2003). Further, adapting to a face at one end of a trajectory through face-space (e.g., the trajectory might have a face with fat lips and thin nose at one end, and a face with thin lips and fat nose at the other) can make the average face look more like the individual at the other end of the trajectory (Leopold, O'Toole, Vetter, & Blanz, 2001). These findings have been interpreted as showing that adaptation to one kind of face can shift the norm of face-space, producing a percept in the opposite direction.

2.3.3 Integrating the notions of face-space and holistic processing.

Face-space is one of the two primary theoretical concepts referred to in explaining behavioural findings with faces; the other is holistic processing. An important question is how these two concepts might relate to each other. There is surprising little theoretical work in this area.

With respect to inversion effects and face-space, Valentine (1991) suggested that inverted faces are rotated to upright before being compared to the faces in face-space. He argued that error is introduced during this process, thus explaining the usual finding of poorer performance for inverted than upright faces. The rotation idea predicts only quantitative not qualitative differences between upright and inverted face processing (Valentine & Bruce, 1988). As this has been shown not to be true (e.g., composite effect, Young et al., 1987), some other explanation must be found. It may be that there is a separate norm for inverted faces, as has been suggested for male faces and female faces (Rhodes et al., 2003). Alternatively, a whole separate face-space might exist for inverted faces. If this were less tightly tuned than the face-space for upright faces, subjects might be worse at discriminating and recognising inverted faces than upright faces. Leder and Bruce (1998) have also suggested that dimensions describing second order relational information (e.g., distance between the eyes) are coded for upright faces but are not coded for inverted faces, while dimensions describing feature shape (e.g., eyebrow thickness) may be coded for both orientations.

Empirically, some studies have combined an inversion manipulation with effects usually explained in terms of face-space. Valentine (1991) found that inversion had a greater effect on typical faces than on distinctive faces. Similarly, inversion effects are larger for other race faces than own race faces (Rhodes, 1993; Rhodes et al., 1989).

With respect to the part-whole effect and face-space, Tanaka et al. (2004) have tested the size of the part-whole effect (Davidoff & Donnelly, 1990; Tanaka & Farah, 1993) as a function of race. Results showed that for German Caucasian subjects (with little experience of Asian faces) there was a large part-whole effect for own-race faces, but no effect for other-race faces. For Asian subjects living in Canada, who had only slightly more experience with Asian than Caucasian faces, there were part-whole effects for both own and other race faces. These results were interpreted as showing that holistic processing is initially weak for unfamiliar face categories, but can develop with contact with the other race, as dimensions of face-space become tuned to better code individuals of that race.

The relationship between second-order relational processing and face-space has also been explored. Thompson et al. (2001) showed that 7-month old infants preferred to look at faces with average eye-mouth distance, than at faces with eye-mouth distances either larger or smaller than average. Conversely, Geldart et al. (1999) showed that 5-month olds preferred faces in which the height of internal features (moved as a group) was higher than average, even though adults found faces with features placed at average height most attractive.

In studies testing distinctiveness effects, Leder and Bruce (1998) increased facial distinctiveness by making either a local feature change (e.g., making the eyebrows thicker) or a spacing/relational change (e.g., moving the eyes closer together). They then showed that for upright faces both spacing and feature changes to distinctiveness increased recognition memory performance compared to memory for the original faces. For inverted faces only feature changes increased memory compared to original. These results were replicated by Gilchrist and McKone (2003) with both adult and 7-year-old subjects, and are consistent with previous findings on featural or spacing changes (e.g., Bartlett & Searcy, 1993; Rhodes, Brake, & Atkinson, 1993).

Finally, the coding of second-order relational information in face-space has been explored by Mckone, Aitkin and Edwards (submitted). They used two relational distortions (changing the height of both eyes together, or changing the height of the two eyes independently) matched in terms of the amount of metric change in the face. In both cases, the eyes could be moved some distance before subjects' perception of normality or identity was affected (e.g., before subjects said that the face no longer looked as normal as the original face). However, the threshold amount by which the eyes could be moved was less for the one-eye-up-one-eye-down manipulation, than for the both-eyes up/down manipulation. This difference in subjects' perception corresponds to variability in the real world (i.e., the range of faces which need to be coded), with more variability found in the position of the two eyes together in the head than in the asymmetry between eyes (see Chapter 6 for details). McKone et al. suggest that their results can be interpreted as the variability in faces experienced affecting the variability of exemplar placement in face-space, which in turn affects subjects' perception. In particular, faces would be clustered more closely together on a one-eye-up-one-eye-down dimension (less variability/smaller coding range), than on a both-eyes up/down dimension (more variability/larger coding range) in face-space.

2.3.4 Face-space and this thesis.

The concept of a face-space, in which individual faces are coded in terms of their position on a variety of underlying dimensions, has proved a useful concept in explaining important empirical phenomena such as distinctiveness, caricature, and cross-race effects. The theoretical relationship between the concepts of face-space and holistic processing, however, remains rather unclear. I suggest that this is partly because very little is known about the possible dimensions of face-space.

In Chapter 6, I use the distortions introduced by McKone et al. (submitted), which correspond to different amounts of variability in face-space to ask whether some dimensions in face-space are more adaptable than others. I also compare adaptation to these distortions for upright faces to that for inverted faces, to see whether there are any differences, and finally I assess whether adaptation transfers between upright and inverted faces as a way of assessing whether different neural populations (and thus face-spaces) are being used for upright and inverted faces.

CHAPTER 3: CAN HOLISTIC PROCESSING BE LEARNED FOR INVERTED FACES?

3.1 Overview.

The origin of "special" processing for upright faces has been a matter of ongoing debate. If it is due to generic expertise, as opposed to being domain specific, holistic processing should be learnable for stimuli other than upright faces. Here I assess this for inverted faces. I trained subjects to discriminate identical twins using up to 1100 exposures to each twin in different poses and images. In the upright orientation, twin discrimination was supported by holistic processing. Removal of a single face feature had no effect on performance, and a composite effect (Young et al., 1987) was obtained. In the inverted orientation, however, above-chance identification ability relied on (a) image specific learning, or (b) tiny local feature differences not noticed in the upright faces. The failure to learn holistic processing for inverted faces indicates that, in contrast to the situation for objects (Tarr & Pinker, 1989), orientation specificity of face processing is highly stable against practice.

3.2 Introduction

For ordinary adults, faces form a special class of visual object. A specific cortical area in the right fusiform gyrus activates more strongly for faces than for other within-class object discrimination (FFA; Kanwisher, McDermott & Chun, 1997), and much evidence shows that a <u>holistic</u> or <u>configural</u> style of cognitive processing makes faces "special". These include the <u>disproportionate inversion effects</u> (e.g., Yin, 1969), <u>composite effect</u> (Young, Hellawell & Hay, 1987), <u>part-whole paradigm</u> (e.g., Tanaka & Farah, 1993) as well as others reviewed in Section 1.2.3.[1]

[1] Examiners are reminded that Experiment 1 has already been examined as my Honours thesis under the title "Does practice induce configural processing of inverted faces?" (2000). It is therefore not eligible for examination, but is included here as Experiments 2 & 3 (which are eligible) follow directly from it. The Experiments in this chapter are published as Robbins & McKone (2003), Cognition, 88(1), 79-107.

3.2.1 Why is holistic processing limited to upright faces?

This chapter compares two general theories of why, in adults, holistic processing is limited to faces, and moreover to upright faces. According to the expertise hypothesis, holistic processing is a domain-general property of expertise in making within-class discriminations. Diamond and Carey (1986) noted that, for most people, faces are the only class of visual stimulus for which they become genuine experts. However, these authors suggested that face-like holistic processing might be learnable for any object type, as long as three conditions were met: (a) all exemplars of the stimulus class share a similar basic configuration (i.e., the same parts in the same left-right, above-below relationships), (b) individual exemplars differ from this shared first-order configuration only in minor (second-order) ways (e.g., the exact distance between two parts; the exact shape of a part), and (c) the subject has sufficient expertise with the stimulus domain to make reliable discrimination between individual exemplars (e.g., dog-show judges who can remember one Scotch Terrier as distinct from another). According to the expertise view, holistic processing for faces is limited to the upright orientation because this is the orientation in which faces are usually experienced.

An alternative theory is that "special" processing for faces in adults is domain-specific, driven by some innate component or by exposure in early infancy (Farah et al., 2000; de Gelder & Rouw, 2001; Morton & Johnson, 1991). This is supported by the finding that young babies orient preferentially towards face-like stimuli (Johnson, et al., 1991), and that there appears to be a critical period in early infancy for the development of holistic processing (Le Grand, Mondloch, Maurer, & Brent, 2001, 2003, 2004). According to this view, the fact that holistic processing develops only for upright faces may be because (a) there is an innate representation of the basic structure of an upright face (eye-like-blobs above nose-like-blob above mouth-like blob), and/or (b) that a bias of subcortical origin in infants' visual orienting causes upright faces to be a frequent input to developing cortical systems (de Hann et al., 2002), and/or (c) very early exposure to upright faces fixes the axes of face-space to suit this orientation. Critically, this view does not deny that experience has a role in face recognition, but proposes that generic perceptual learning in adults reflects different mechanisms from those involved in learning upright faces in infancy.

3.2.2 Predictions of a generic expertise account.

If holistic processing is due to generic expertise then, as an adult, it should be possible to learn holistic processing for stimuli other than upright faces. In the literature to date several studies have explored this prediction via tests on experts with non-face objects (a more detailed review of this literature is given in Chapter 5).

For objects-of-expertise, disproportionate inversion effects have been obtained in long term memory with dog experts (Diamond & Carey, 1986), although these have not been replicated in a sequential matching task with car and bird experts (Gauthier, Skudlarski, Gore, & Anderson, 2000). Relatively few studies have used paradigms which directly assess holistic processing. Using the part-whole paradigm, Tanaka et al. (cited in Tanaka & Gauthier, 1997) found some suggestion of an expertise effect – that is, a larger whole-part advantage in experts than in novices – for dog experts (interestingly, looking at dog faces), but not for car experts or biological cell experts. Other studies have investigated experiment-trained "greeble experts". Greebles are an artificial object class that may be grouped into genders (by direction of protrusion) and families (by body shape); experts are trained using seven to ten 1-hr sessions involving identification of greebles at the gender, family and individual-name levels. Three greeble studies have shown no reliable training effect in the basic part-whole paradigm (Gauthier & Tarr, 1997; Gauthier & Tarr, 2002; Gauthier et al., 1998), and no composite effect in trained subjects (Gauthier & Tarr, 2002; Gauthier et al., 1998). These studies have, however, shown some suggestion of a face-like effect in a modification of the part-whole paradigm: one of three greeble parts produced better memory for the part in the original configuration than in an altered configuration (cf. Tanaka & Sengco, 1997). Overall, studies with object experts have produced some suggestion of holistic processing, although the evidence is currently less than convincing.

3.2.3 Holistic processing for inverted faces?

In addition to non-face objects, another stimulus class of interest is inverted faces. If it were the case that, as an adult, holistic processing could be learned for inverted faces, then this would provide compelling evidence for the expertise hypothesis.

An important question is then how much practice would be required. Two sets of literature are relevant to this issue. First, the authors of the greeble studies have suggested that expertise sufficient to support holistic processing can be learned in under 10 hours of practice. Second, results from speeded object-naming tasks indicate that, for non-face objects, processing in the inverted orientation can come to share properties exhibited in the upright orientation with only a very small amount of training.

When objects are rotated in the picture plane they are initially named more slowly the further they are from upright (e.g., Jolicoeur, 1985), but this effect disappears with practice (for review, see McKone & Grenfell, 1999). For objects requiring within-class discrimination, Tarr and Pinker (1989) showed that, after practice at specified new orientations, naming times at intermediate positions then increased as a function of distance from the nearest-trained-orientation. They interpreted these results as evidence for "view-based" (i.e., template-like) representations of objects regardless of orientation: prior to the experiment, most stored views of a familiar object are in the upright (canonical) orientation and, within the experiment, new views are rapidly formed following exposure to novel orientations. Importantly in the present context, these upright-like representations in new orientations were created in less than 100 trials.

No previous studies have adequately assessed whether it is possible to learn holistic processing of faces in the inverted orientation. Occasional claims that holistic processing can be learned for inverted faces are partly a result of miscitation. Both Valentine (1988) and Sergent (1984) cite Bradshaw and Wallace (1971) as finding no difference between upright and inverted faces following practice; however, with the particular task used in that study, Bradshaw and Wallace in fact reported that both orientations were processed in a part-based manner. Valentine (1988) also contrasted the findings of Sergent (1984), who found only part-based processing for inverted faces in unpracticed subjects, with those of Takane and Sergent (1983), who he claimed found holistic processing after practice; however, Takane and Sergent did not actually present or analyse any data for their inverted condition (although they did state that it was "similar to the upright condition", p. 405). Endo, Masame and Maruyama (1990) provide the only direct claim of holistic processing in inverted faces after practice. They used highly schematic faces (e.g., an unusual head outline; circles for eyes; triangle for nose etc.) in a vertical half-face version of the composite paradigm. The standard pattern – a composite effect upright but not inverted – was obtained when subjects were unpracticed. Following extensive training with inverted faces a composite effect did emerge in a condition with different headshapes in each half face. I suspect, however, that this could be attributed to the extreme violation of vertical symmetry that resulted in the aligned condition. When the two halves of the head were symmetric, there was no composite effect for inverted faces even after training.

In contrast to this suggested evidence for holistic processing of inverted faces, another series of studies (Martini, McKone, & Nakayama, 2002, cited in McKone, Martini, & Nakayama, 2003; McKone, 2004; McKone, Martini, & Nakayama, 2001) argue against any such learning. Each of these studies was designed to isolate the holistic component of face processing, by identifying some phenomenon which existed for upright whole faces, but was completely absent for inverted faces. In the present context, the relevant point is that subjects were given hundreds or thousands of trials with the face stimuli in the inverted orientation, and yet showed no signs of developing the signature phenomena for holistic processing. Only a limited style of practice was used, however, presenting the same image repeatedly, rather than different views, as in real life. Farah et al. (2000) note that recognising faces across different views is something prosopagnosics cannot do, suggesting that it requires holistic processing. Similarly, Tong and Nakayama (1999) state that a variety of views and contexts are needed to acquire a "robust representation" of a face. Thus, seeing an inverted face over a variety of views may be necessary to acquire a full holistic representation.

3.2.4 Present Study

The aim of the present study was to assess whether, with appropriate practice, inverted faces could come to be processed holistically. A major aspect of the design was the use of identical twins as stimuli to encourage maximum reliance on holistic rather than part-based processing. In real-world face recognition, single local features do not generally differentiate people reliably (e.g., many individuals have blue eyes). In an experimental setting, however, where stimuli include a limited number of different faces, local features can contribute substantially to performance. Even discrimination of approximately similar individuals (e.g., the same sex and age) could be based on local information alone (e.g., eye colour; presence of a particular freckle), especially when subjects see the same faces over many hours of practice.

Thus, to give the best chance for any holistic processing for inverted faces to emerge, I wished to minimise local feature cues that might be used to identify the faces.

The hope was that, with identical twins, no single feature would differ enough between siblings to support reliable discrimination, and instead that identification would rely on information integrated across the entire face region (i.e., holistic processing). Use of multiple images and viewing angles also made very local information (e.g., exact shape at the corner in the mouth in one particular photograph) unreliable as a cue to identity, and made learning more similar to real-life (see discussion on multiple views above).

During training, each twin (e.g., "Liz Smith") was individually named approximately 350 times (Experiment 1) or 280 times (Experiment 2). This level of practice exceeded the level of practice used in the greeble studies of trained expertise (120 training trials per individual greeble; Gauthier & Tarr, 1997), and also the number of trials in the "naming rotated objects studies" necessary to produce upright-like representations at other orientations (less than 100 trials per object; Tarr & Pinker, 1989). Thus, although a training study can never hope to match the degree of real-world experience that people have with upright faces (or, for example, that which expert dog-show judges have with their breed of expertise), I argue that the present study will at least provide a strong answer to the question of whether holistic processing for inverted faces can, or cannot, emerge in experiment-trained "experts".

3.3 Experiment 1 – Learning to recognise twins

In Experiment 1, subjects were trained to identify two sets of female identical twins, given pseudonyms Liz and Ruth Smith, and Ann and Clare Brown. Orientation was a between-subjects variable. Each subject completed 8 hrs of training sessions, in which feedback was given for decisions at three levels of categorisation (cf. Gauthier & Tarr, 1997), namely the individual level (e.g., Is this Liz?), the family level (e.g., Is this one of the Smiths?), and a gender level (e.g., Is this a female?). Individual level trials were of primary interest. The family level provided a check on subjects' ability to learn arbitrary face-name associations. Note that subjects should show good performance at both the gender and family levels, even in the inverted orientation, since local features (e.g., the lighter eyes of one set of twins) would be sufficient to support decisions at these levels.

Experiment 1 included several tasks. Following the training phase (Experiment 1a), all subjects participated in tests designed to ascertain what, if anything, had been learned. First, subjects gave verbal strategy reports (Experiment 1b). Subjects were then given a generalisation of identification test (Experiment 1c), which examined ability to identify new photographs of the twins. This assessed whether learning was based on particular training images. The task was performed both with eyebrows and without eyebrows visible, since the strategy reports had indicated that this was the feature which differed most between the twins. Finally, a second test of generalisation was a same-different task (Experiment 1d), again with new photographs, and again both with and without eyebrows.

I predicted that subjects in the upright face condition would: (a) learn to identify twins accurately at the individual level in the training phase (cf. 110 trials to reach 90% accuracy for naming twins in Stevenage, 1998); (b) generalise this knowledge to previously unseen photographs; (c) differentiate same-twin pairs (Liz-Liz) from different-twin pairs (Liz-Ruth) with relative ease; and (d) remain unaffected by the removal of the eyebrows. This last test provides a direct assessment of holistic processing. Previous evidence argues that holistic processing is not substantially impaired by removing a single face feature: for example, accuracy of naming famous faces is unaffected by removing only the eyes, or nose, or mouth (Experiment 17, Moscovitch et al., 1997).

For subjects in the inverted face condition, several outcomes were considered possible. A pattern similar to that just described for upright faces would suggest that holistic processing had emerged with practice. Alternatively, failure to develop holistic processing could be reflected in several ways. First, given the difficulty of discriminating twins based on local information, subjects might fail to learn the twins at all during training. Second, learning might be good during training, but fail to generalise to new photographs. Third, learning might be good, but (despite my best efforts with the stimuli) be shown to rely solely on local features.

3.3.1 Experiment 1: General Method

3.3.1.1 Subjects and Overall Design.

Subjects were aged 18-23 yrs, were of Caucasian background (i.e., the same race as the stimulus faces) and reported normal or corrected-to-normal vision. Trained subjects were friends of mine and were not paid. Each subject saw faces only in one orientation in all phases of the experiment, with N=9 in the inverted condition, and N=6 in the upright condition.

Trained subjects participated in ten 1-hr sessions, spread over approximately 5 weeks. Sessions 1-8 comprised training with feedback, during which each of the twins was seen approximately 1100 times (in total). Session 9 tested generalisation of identification to new photographs, and the same-different task, using stimuli with the eyebrows shown as normal. Session 10 was identical to Session 9, but this time eyebrows were covered on all stimuli. Subjects were asked not to communicate with any other participant regarding the study until the completion of the experiment.

Total testing time for Experiment 1 was 181 hours. This includes 150 hrs for the 15 trained subjects, plus 31 hours to test control subjects for the same-different task (see Experiment 1d).

3.1.1.2 Face Photographs.

Sample stimuli for each twin are shown in Figure 3.1a. Two sets of Caucasian female identical twins (aged 22 years and 15 years) were photographed from seven different views: full-front (0°) and –20°, –15°, –5°, +5°, +15° and +20° from full-front (see Figure 3.1b). There were 138 neutral expression greyscale photographs of each twin (16 at each angle plus an extra 26 full-front). Photographs were cropped to remove external features and digitally retouched. For the Smith twins this included an attempt to thicken Liz's eyebrows, which were plucked differently to Ruth's. Each stimulus measured 55 by 37 mm at a screen resolution of 1024 by 768 pixels. All manipulations were performed using Photoshop 5.0.

Additionally, 36 cropped photographs of Caucasian male faces (13 males each pictured at –25°, +25° and full-face; see Figure 3.1c) were used in gender decisions.

These came from the Harvard Vision Lab face database (F. Tong & K. Nakayama). Brightness and contrast were equated across the male and female sets, as well as for all four twins.

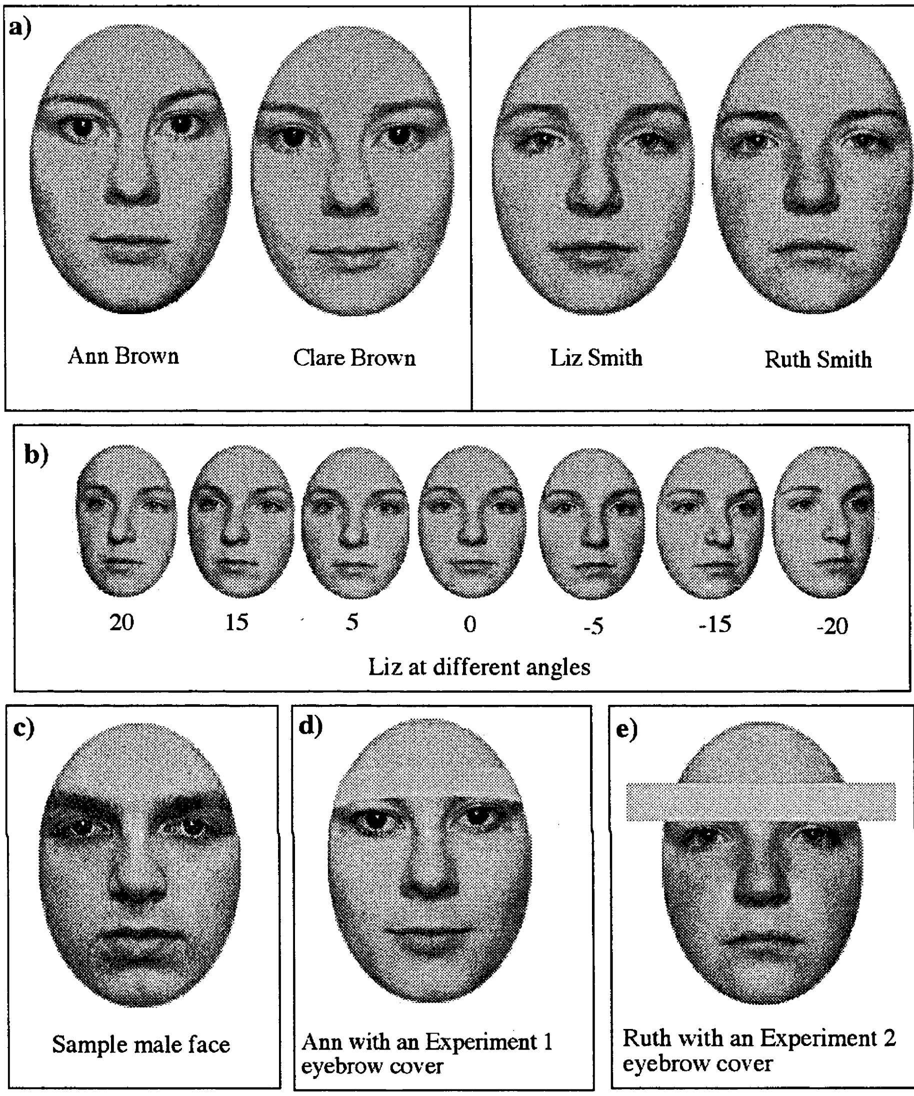


Figure 3.1. Sample stimulus photographs.

3.1.2 Experiment 1a: Identification training.

Training involved categorisation of twins at three levels: gender, family and individual. It was predicted that sufficient local featural information would exist to distinguish gender and family even in the inverted condition. It was also predicted that, upright, subjects would be able to learn to identify the twins at the individual level; the first question of interest was whether subjects in the inverted condition could do so.

3.1.2.1 Method.

Selection of training phase photographs. Each subject's training set of 69 photographs of each twin (276 total) was randomly selected from a general set of 91 photographs of each twin[2]. Similarly 20 of the 36 male photographs were selected for each subject's training set. The remaining 22 photos of each twin (and 16 of males) were held aside as the generalisation set (Experiment 1c) for that subject.

Each training session contained a total of 296 trials (40 gender, 80 family and 176 individual). Within each session, photographs from each subject's training set were randomly assigned to questions (half of which had a "yes" answer and half "no").

Procedure. Each trial comprised: a yes/no question based on either the gender, family or individual name (800 ms), a blank screen (100 ms), a fixation point (300 ms), a face until the observer responded (up to 3,000 ms), another blank screen (1,500 ms), and finally the same face with the correct answer shown below it (3,000 ms). The inter-trial interval was 200 ms. Subjects were given a self-timed rest every 50 questions. Prior to testing, subjects were told there were two sets of identical twins but not given any names, forcing them to guess on early trials. Subjects were initially instructed to concentrate on accuracy, but were told to also aim for speed when they reached an accuracy of 90% at all levels.

Stimuli were displayed using PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993) on a Macintosh Power PC 7200/75. Subjects sat approximately 50 cm from the screen, giving visual angles for the faces of 6.3° vertical by 4.2° horizontal.

3.1.1.2 Results and Discussion.

Table 3.1 shows mean identification accuracy during the training phase (Sessions 1-8). For gender decisions, subjects in both orientation conditions were, as expected, highly accurate even in Session 1.

---

[2] The other 47 photographs were used for other tasks.

For family decisions, subjects quickly learned face-name associations, with performance approaching ceiling for both orientations in Session 2. Learning was slower at the individual level than at the gender or family levels, but average accuracy was reasonably good within the first two or three sessions, and reached over 75% in Sessions 7 and 8. Accuracy for the Smith twins was somewhat lower than for the Brown twins, and there was a small but consistent inversion effect [3].

Table 3.1. Experiment 1a: Average accuracy (% correct) over training sessions for the gender, family and individual categorisations (the Smiths and Browns are shown separately).

	Sess 1	Sess 2	Sess 3	Sess 4	Sess 5	Sess 6	Sess 7	Sess 8
Gender								
Upright	98	98	99	99	99	99	99	98
Inverted	95	99	98	99	98	98	99	98
Family								
Upright	85	98	98	95	95	97	98	97
Inverted	76	93	95	96	95	94	96	97
Individual (Browns)								
Upright	64	79	87	90	93	92	96	95
Inverted	63	77	83	86	86	87	90	91
Individual (Smiths)								
Upright	62	71	77	79	75	82	82	82
Inverted	52	64	70	71	70	72	77	78

Note: Chance = 50%

Further analysis focused on individual subject data shown in Figures 3.2, 3.3, and 3.4 for all tasks of Experiment 1. The first column shows identification accuracy (% correct) in the final two training sessions (Sessions 7 and 8, combined to give more power). To assess whether or not every subject learned to discriminate the twins, end of

[3] Reaction times (RTs) were also measured. For upright faces, mean RTs (correct answers only) across the 8 sessions were: Gender = 921 (Session 1), 851, 717, 601, 591, 651, 638, 551 ms (Session 8); Family = 1250, 1073, 884, 826, 716, 762, 690, 664; and Individual = 1401, 1334, 1150, 961, 837, 928, 814, 789. Interestingly, the RTs for the three levels do not appear to be approaching the same asymptote with practice. In sessions 7 and 8, an overall effect of level remained, $F(2,10) = 9.18$, $MSE=7119.2$ $p<.01$, with RTs to identify individuals 207 ms slower than gender decisions, $F(1,10) = 18.1$, $p<.05$, and 125 ms slower than family decisions, $F(1,10) = 6.57$, $p<.05$. This result for upright faces (for which subjects are by definition experts) contrasts with Gauthier and Tarr's (1997) use of equal RTs at the three levels as a criterion for expertise with greebles.

training performance was compared with chance using the binomial distribution. With 176 trials per family, a one-tailed comparison requires an identification accuracy of greater than 57% to be significantly better than chance (50%), at $p < .05$.

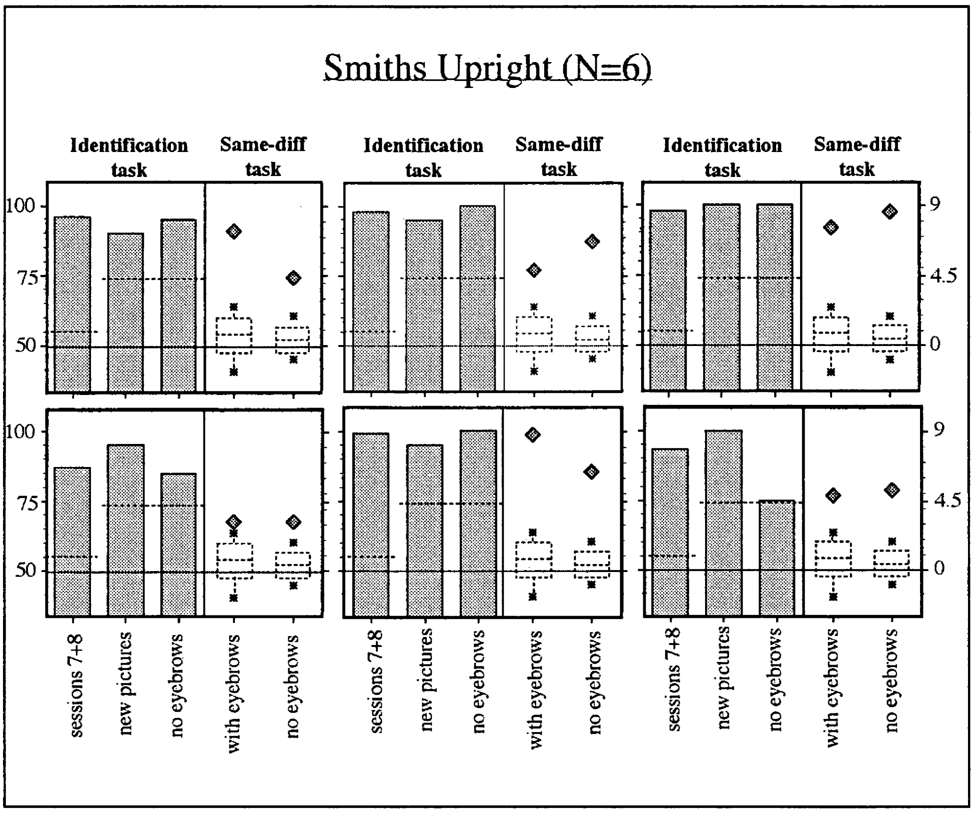


Figure 3.2. Experiment 1: Individual subject results in the upright condition, on the Smith twins. For the identification task, the leftmost column shows accuracy (% correct) in training sessions 7+8 (Experiment 1a), the next two columns show the generalisation tests both with and without eyebrows visible (Experiment 1c), and the dotted lines show the accuracy needed to be significantly better than chance. For the same-different task diamonds indicate the trained subject's discriminability score with and without eyebrows (Experiment 1d). Control subject scores are shown as box plots indicating mean, ± 1 SD and range.

For the Smith twins, in the upright orientation (left-most bar of Figure 3.2), all subjects learned to identify the twins accurately by the end of training. In the inverted condition (left-most bar of Figure 3.3), eight of nine subjects also identified the Smiths at well above chance levels; Subject 12, however, failed to learn the Smiths, the reason for which will become apparent when further tests are reported. For the Brown twins (left-most columns of Figure 3.4a, upright; and Figures 3.4b and 3.4c, inverted), all subjects were well above chance in sessions 7 and 8, for both orientations (note that subjects in the inverted condition are grouped on a critical difference which emerged in their behaviour on later tests).

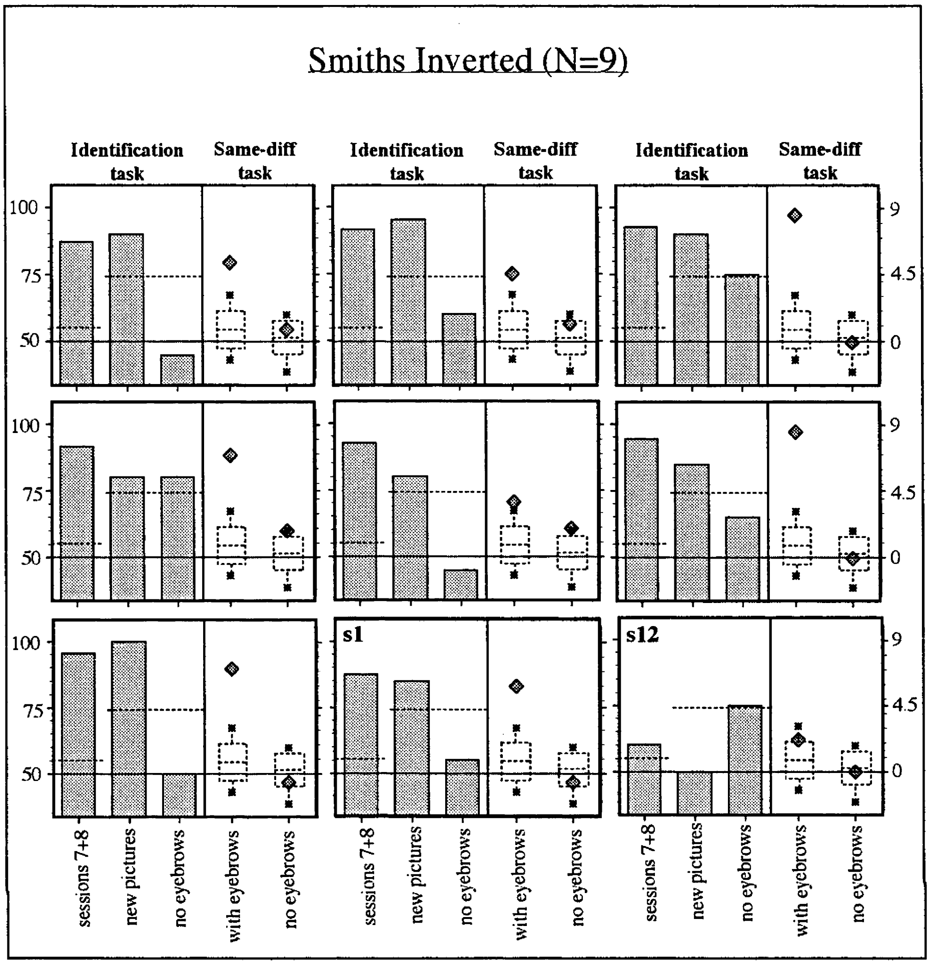


Figure 3.3. Experiment 1: Individual subject results in the inverted condition, on the Smith twins. Figure formatted as in Figure 3.2. Two subjects of particular interest are numbered (S1 & S12).

In summary, Experiment 1a showed that all subjects learned to identify the individual Brown twins reliably. The Smiths were more difficult to distinguish, but all subjects in the upright condition learned them well, and most of the subjects in the inverted condition also did so eventually. It was then necessary to ascertain exactly what had been learned: individual photographs, local features, or a true holistic representation.

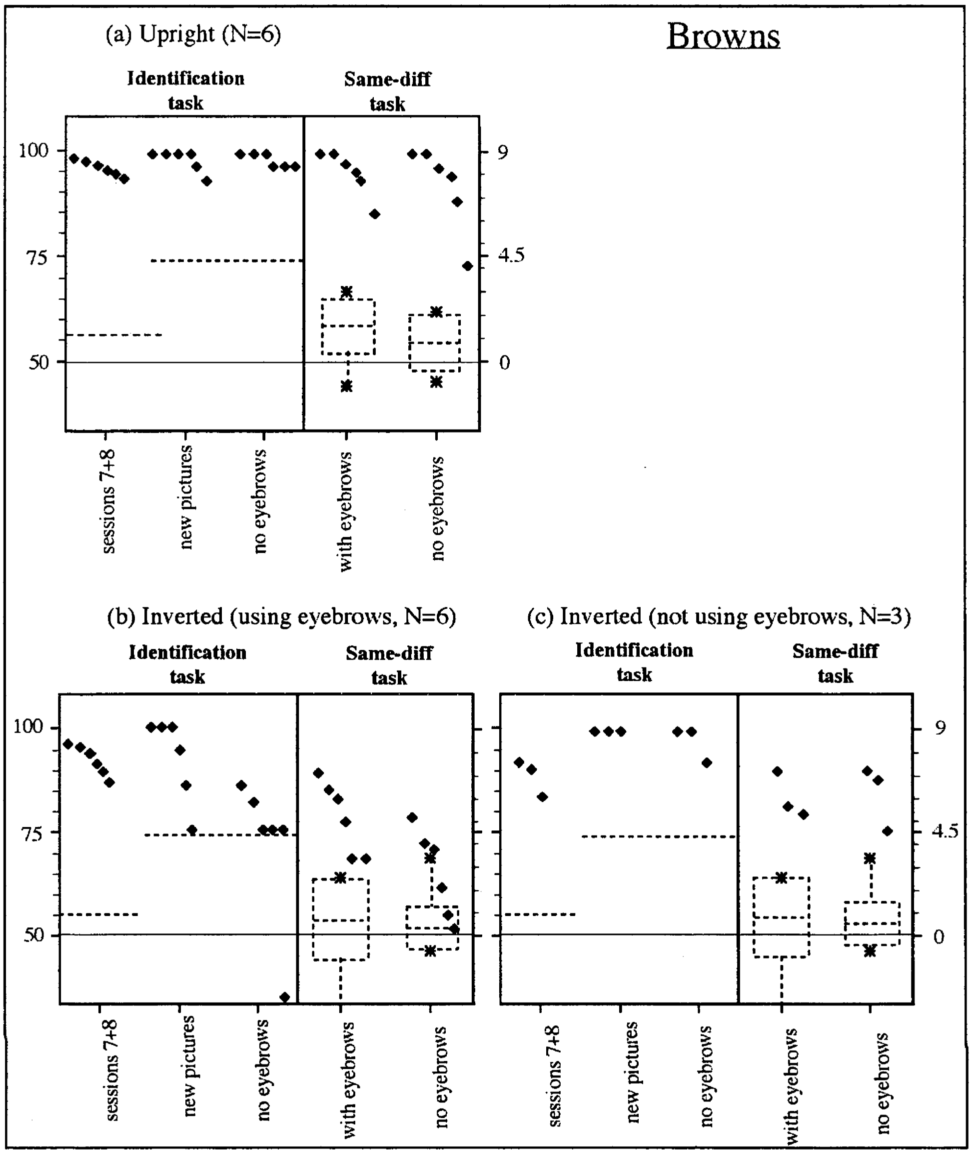


Figure 3.4. Experiment 1: Results for the Brown twins, for (a) the upright condition, (b) those in the inverted condition who reported using eyebrows to differentiate the twins and (c) those in the inverted condition who did not report using eyebrows. To save space each plot contains multiple subjects showing a common pattern (each diamond represents one trained subject), otherwise formatting is as in Figure 3.2.

3.1.3 Experiment 1b: Self reported strategies.

Spontaneous comments from subjects in the inverted condition suggested that, during training, many were focussing on a single local feature – eyebrow grooming – to differentiate the twins. (Even after my touching up of the original photographs, there remained small differences between the eyebrows of each sister; these are best seen in Figure 3.1a if the faces are turned upside-down). To investigate the role of the "eyebrow" strategy further, subjects were formally asked to list their strategies.

Two independent judges, blind to subjects' condition, rated strategy reports in terms of whether these mentioned eyebrows. Inter-rater reliability was high (Cohen's kappa = 96.6%). Example eyebrow strategies included "Ruth's eyebrows have little thin bits in the middle while Liz's don't" (Smiths) or "Right eyebrow: Ann has a diagonal slash missing" (Browns). I did not attempt to code holistic/configural strategies, since this type of information is not well described verbally (e.g., Fallshore & Schooler, 1995).

Table 3.2 shows the number of subjects using a local "eyebrow" strategy. In the upright condition, no subject used eyebrows for individual-level decisions, for either twin pair. In the inverted condition, however, most subjects used eyebrows for telling apart individual twins. This was most consistent for the Smiths, where the only subject who did not mention eyebrows was Subject 12; that is, the person who failed to learn to differentiate these twins during training. For the Browns, 6 subjects (of 9) used eyebrow differences, with other strategies reported also tending to be local in nature.

Table 3.2. Experiment 1b: Number (and percentage) of subjects in the upright and inverted conditions rated as using eyebrow cues for gender, family, and individual categorisations.

	N	gender	family	Browns	Smiths
Upright	6	2 (33%)	0 (0%)	0 (0%)	0 (0%)
Inverted	9	7 (78%)	1 (11%)	6 (67%)	8 (89%)

Thus, for the Smith twins, strategy reports suggest that the sisters were so similar to each other that subjects in the inverted condition could learn to tell them apart only if they noticed the eyebrow difference. In addition to Subject 12, further support for this idea comes from Subject 1: during training this person failed to learn the Smiths for the first 5 sessions, but then said that he had finally "cracked it!" and it was "in the eyebrows" after which his performance dramatically improved. I presume that the broader range of local strategies reported for the Brown twins reflects the fact that these twins have more differences between them than the Smiths (see Figure 3.1a).

3.1.4 Experiment 1c: Generalisation in the identification task

This part of the experiment was partly designed to assess whether learning in the training phase reflected learning of particular training images of each twin, by testing generalisation of identification accuracy to previously unseen photographs. In addition, the task was run both with eyebrows visible (Session 9) and without eyebrows visible (Session 10; using a "headband" see Figure 3.1d), to assess whether successful identification performance relied on the presence of the local feature most commonly mentioned in the strategy reports.

3.1.4.1 Method.

The basic design and procedure were the same as in the training phase (i.e., confirm the identity of each twin presented one-at-a-time). A "name reminder" block (48 trials) used training photographs with feedback. A pre-generalisation block (48 trials) then used training photographs but familiarised subjects with a no-feedback procedure. This merged seamlessly into the generalisation block using previously unseen photographs without feedback. The generalisation block included 104 trials; comprising 32 gender level, 32 family level, and 40 individual level trials (10 of each twin).

3.1.4.2 Results and Discussion.

Generalisation to new photographs at the gender and family levels was excellent (> 90%) as expected, so only results for the individual level will be discussed. In Figures 3.2 through 3.4, the second-from-left column indicates identification accuracy for new photographs with eyebrows included, and the third-from-left column indicates accuracy without eyebrows visible. For a significant difference from chance, with 20 trials per family, a one-tailed comparison requires an identification accuracy of greater than 74%, at $p < .05$. Table 3.3 shows further statistical analyses contrasting, for each subject, the identification accuracy at the end-of-training (Sessions 7+8) with the accuracy (a) in the generalisation test with eyebrows, and (b) in the generalisation test without eyebrows. Results of chi-square tests for each subject indicate the number who showed a significant decrement between training and generalisation.

In the upright condition, all subjects generalised well to new photographs that included eyebrows (see Figure 3.2, Figure 3.4a, and Table 3.3), with identification accuracy remaining high, and no significant decrement compared to end-of-training for both sets of twins. When the eyebrows were covered, upright performance generally remained well above chance (Figure 3.2 and Figure 3.4a); moreover, no subject showed a significant decrement in comparison to the end of training for the Smiths, and only one did so for the Browns (Table 3.3). These results indicate that subjects in the upright condition had not merely learned particular photographs used in the training phase. In addition, removal of a single local feature from the image had essentially no effect on identification accuracy, consistent with the use of holistic processing (cf. Moscovitch et al., 1997).

In the inverted orientation for the Smiths (Figure 3.3), all except Subject 12 were above chance for new photographs that included eyebrows, and no subject showed a significant decrement in accuracy as compared to the end-of-training (Table 3.3). Thus, identification generalised well, arguing against learning of individual images. Critically, however, identification accuracy for the Smiths dropped substantially when the eyebrows were covered. Excluding Subject 12, performance in the without-eyebrows condition dropped to below-chance levels for 6 of the 8 subjects, with a seventh only just above chance; the drop was a significant change from end-of-training in all 7 of these cases. Thus, as suggested by the self-reported strategies, subjects trained on the inverted Smiths were relying substantially on eyebrow differences between Liz and Ruth, and had been able to learn very little distinguishing information about the rest of the Smiths' faces.

For the inverted Browns, recall that local feature strategies other than the eyebrows were reported by some subjects; thus, I would not expect those subjects to be affected by only covering the eyebrows, and hence results are shown separately for those in the inverted condition who reported using eyebrows (N=6; Figure 3.4b), and those who did not (N=3; Figure 3.4c). When eyebrows were intact all but one subject generalised well to new photographs (Figure 3.4b and 3.4c, and Table 3.3). When eyebrows were covered, however, subjects who reported an eyebrow strategy showed a drop in performance (Figure 3.4b) which was significant in 3 of 6 cases (Table 3.3), while subjects who did not report using eyebrows were substantially less affected (Figure 3.4c; Table 3.3).

Table 3.3. Experiment 1c: Summary of chi-square tests comparing identification accuracy for new photographs, both with and without eyebrows, to the end of training ($\chi^2_{crit}(1, N=196) = 3.84, p < .05$). Data are number of subjects who showed a significant drop in performance.

	Smiths	
Subject group	With eyebrows	Without eyebrows
Upright	0 (of 6)	0 (of 6)
Inverted	0 (of 8*)	7 (of 8*)
	Browns	
	With eyebrows	Without eyebrows
Upright	0 (of 6)	1 (of 6)
Inverted (using eyebrows)	1 (of 6)	3 (of 6)
Inverted (not using eyebrows)	0 (of 3)	0 (of 3)

Notes: * Excluding Subject 12, who could not identify the Smiths at the end of training.

In summary, the results of this first generalisation test indicate that subjects had not merely learned individual photographs in the training phase, in either orientation. This did not mean, however, that a holistic representation had been learned by subjects trained on the inverted faces. Instead, most of these subjects had merely learned a representation of eyebrow differences, particularly for the Smiths.

3.1.5 Experiment 1d: Generalisation to a same-different task

The second test of generalisation used a new task to assess subjects' ability to distinguish the twins. In this same-different task, a pair of photographs were shown together, and subjects indicated whether these were of the same person (e.g., Liz-Liz), or of a twin and her sister (e.g., Liz-Ruth). Another set of new photographs was employed, and the task was performed both with eyebrows (Session 9) and without eyebrows (Session 10).

The same-different task was included because I thought that it might be more sensitive than the simple identification test in showing any learning of inverted faces. In particular, it was possible that subjects might have learnt enough about the twins to be able to tell them apart when seen simultaneously, even though the representation in memory was not good enough to identify each twin reliably when seen one at a time.

3.1.5.1 Method.

Control subjects. To assess the "chance" range of performance control subjects, who had no identification training on the twins were tested. It was expected (cf. Stevenage, 1998) that controls would find it essentially impossible to tell the twins apart. There were 26 untrained subjects (13 upright and 13 inverted) who performed the same-different task with eyebrows visible, and 36 (17 upright and 19 inverted) who performed it without eyebrows visible. Control subjects were first year psychology students, reported normal or corrected to normal vision, were of Caucasian background, and received course credit for the half-hour experiment.

Design. Trained and untrained subjects were shown two types of photograph pairs: either two different photographs of the same person (a same-twin pair) or a photograph of the person and one of their twin (a different-twin pair). Only the five internal views were used (-15°, -5°, full-face, 5°, 15°), and pairs were always formed across different views (e.g., -15° and 5°) so that similarity would be judged on identity rather than pose. A measure of perceived similarity, on a ten point scale (after Stevenage, 1998), was derived from a forced choice decision of whether the pair was same or different identity, plus a confidence rating on a five-point scale (where 1 was totally guessing and 5 was completely sure). The conversion procedure is shown in Table 3.4. This combined accuracy-confidence score produced the same pattern of results as accuracy scores alone, but was used as the dependent variable because it provided a more sensitive measure. For trained subjects the task was performed with eyebrows visible in Session 9 and without eyebrows in Session 10.

Stimuli. Photographs for Experiment 1d were selected by a pilot study (N=12) in which subjects rated 20 same-twin and 20 different-twin pairs for each set of twins (80 pairs in total). From these 80 pairs, 10 same and 10 different pairs for each set of twins were chosen so that, as far as possible, pilot similarity scores were equated across same and different sets. To do so it was necessary to select those same pairs judged most dissimilar to each other, and those different pairs judged most similar to each other. The final stimulus set was therefore 10 same and 10 different Smith twin pairs, 10 same and 10 different Brown twin pairs[4]. Pictures were of the same dimensions as in the training and generalisation tasks.

Table 3.4. Experiment 1d: Conversion of same-different responses and confidence ratings to a ten-point similarity scale.

Response	Confidence	Similarity Score
different	5	1
different	1	5
same	1	6
same	5	10

Procedure. All subjects were told they would see two sets of identical twins, and that they were to decide whether pairs showed the same person or different people. Faces were displayed in pairs with approximately 50 mm between them on the screen. Each pair was displayed for 3,000 ms, allowing for approximately two fixations per face. This was followed by a prompt for the forced choice same-different decision, and then one for a confidence rating on the five-point scale.

3.1.5.2 Results and Discussion.

A discriminability score for each pair of twins was calculated for each subject as their mean different-twin similarity rating minus their mean same-twin similarity rating. This gave a discriminability score of 0 if no difference was perceived between same- and different-twin pairs, and a discriminability score of 9 for maximum perceived difference (i.e., the subject got all decisions correct with the highest level of confidence). A sample calculation is shown in Table 3.5.

[4] One of these photographs of Ann was inadvertently used in training; thus scores were calculated only from the 9 previously unseen different Brown twin pairs.

Discriminability scores for individual subjects are shown in the "same-different task" panels of Figures 3.2-3.4. The control subject scores are indicated as box-plots on each graph. As expected, untrained subjects could barely tell the twins apart: the mean discriminability score for controls was close to zero both with and without eyebrows visible.

Table 3.5. Experiment 1d: Sample calculation of discriminability scores for one subject (upright with eyebrows).

	Smiths		
	Similarity rating Same pairs	Similarity rating Different pairs	Discriminability score Different - Same
Subject 2	2.40	9.80	7.40

Turning to the results for the trained subjects, different patterns emerged for the upright and inverted conditions. In the upright orientation, performance with-eyebrows was good for both the Smiths (see Figure 3.2, fourth column from the left) and the Browns (Figure 3.4a); that is, all trained subjects' discriminability scores were above even the best of the controls. Moreover, discriminability remained good even without-eyebrows, again for both the Smiths (right-most column in Figure 3.2), and the Browns (Figure 3.4a).

In the inverted orientation, for the Smith twins, discriminability with-eyebrows visible was generally good (Figure 3.3, fourth column from the left), excluding Subject 12 (the person who never learned these twins in training). Without-eyebrows visible (Figure 3.3, right-most column), however, performance dropped dramatically, with no trained subject outside the range of controls, and the average across all trained subjects being approximately zero. For the Brown twins, all subjects showed good discriminability with-eyebrows visible (Figures 3.4b and 3.4c, fourth column from the left). Without-eyebrows visible, subjects who reported using eyebrows as a strategy showed a drop in performance (Figure 3.4b, right-most column), while those who did not mention eyebrows in their strategy reports were unaffected (Figure 3.4c, right-most column).

In summary, subjects in the upright condition showed good generalisation to the same-different task (i.e., a new task using new photographs), and once again showed no sensitivity to the removal of eyebrows. This is consistent with the assumption that, upright, a holistic representation of each twin had been learned. In contrast, subjects in the inverted condition were again relying primarily (and in the case of the Smith twins, entirely) on eyebrow information. Thus, these subjects had not picked up enough information about the rest of the twins' faces to tell them apart even when they were seen simultaneously.

3.1.6 Experiment 1 - Summary

The results of Experiments 1a-1d, taken together, lead to a straightforward conclusion. As expected, subjects who were trained on upright faces learned the twins very well, generalised this knowledge to new photographs and a new task, and showed no noticeable decrement in performance when the most discriminating local feature (the eyebrows) was covered from view. Given that holistic processing is well-known to occur for upright faces, I presume that this pattern reflects subjects having learnt a whole-face representation of each individual twin.

In contrast, subjects who learned the twins in the inverted orientation showed evidence of using only local feature strategies, even after 1100 exposures to each twin during the training phase. In the case of the Smith twins, this finding was particularly striking: all subjects used the same feature (a minor difference in eyebrow combing/plucking) to differentiate the twins, and when this feature was covered, subjects could no longer tell the Smiths apart. Thus, Experiment 1 strongly suggests that practice with inverted faces does not induce holistic processing. This is consistent with the failure to develop holistic processing for inverted faces in several earlier studies (Martini et al., 2002, cited in McKone et al, 2003; McKone, 2004; McKone et al. 2001). However, while those studies used extensive practice with a single image, the present study provided experience with many different images of each twin in a variety of views, confirming that even these broader learning conditions do not lead to development of a holistic representation for inverted faces.

It is also worthwhile emphasising that the local feature selected by the inverted subjects is hardly an obvious one. An “eyebrow” strategy is not likely to support real-world or generalisable differentiation between twins: differences in eyebrow plucking can only be seen on close inspection, and, in the real world, people can change their exact patterns of eyebrow combing from day to day. Moreover, most of the subjects in the upright condition did not even notice the eyebrow differences. Thus, it is not the case that the current stimuli included a striking featural difference that might have outweighed subjects' attempts to learn the twins' faces holistically. Instead, for the Smith twins at least, it seems that the faces were so similar that subjects were searching desperately to find any way of distinguishing between them (this was certainly the case for Subjects 1 and 12).

3.2 Experiment 2 - Learning without eyebrows

In Experiment 1, subjects failed to learn a holistic representation of the inverted faces, despite extensive practice. In Experiment 2, I explored what would happen if the featural cue used by subjects in the first experiment was unavailable. Specifically, only the Smith twins were used, and new subjects were trained with the twins' eyebrows covered at all times.

With the most discriminating feature removed from the Smith's faces, several outcomes were considered possible. First, subjects trained on inverted faces might fail to learn the twins at all (cf. Subject 12 in Experiment 1). Second, subjects might learn somewhat, but fail to generalise to new photographs and task, suggesting they had been forced to rely on remembering individual photographs. Third, subjects might learn and generalise well, but have done so only by finding some new local feature to discriminate the Smiths. The reason for running Experiment 2, however, was to test another alternative, namely that removal of the easiest-to-find feature cue might force subjects to learn a holistic representation.

3.2.1 Experiment 2 - Method

3.2.1.1 Subjects and Design.

New subjects were selected for training in Experiment 2. All were university students, paid $10 per session or first year psychology students participating for course credit. There were N=9 in the inverted face condition (age 19-29) and N=8 in the upright face condition (age 18-41). As in Experiment 1 all were Caucasian and reported normal or corrected-to-normal vision.

The amount of identification training for each twin remained similar in Experiment 2 (288 individually-named training trials per twin) to that in Experiment 1 (352 individually-named training trials per twin). This was achieved in only four sessions by dropping the gender level of categorisation training, and including only the same-different task as a test of generalisation.

3.2.1.2 Stimuli and Procedure.

Unless otherwise specified, the procedure for each task was as for Experiment 1. For the training phase, stimuli included 92 photographs of each twin taken from Experiment 1a. Each full training session (Sessions 1-3) included five subblocks of 56 trials, with Session 4 including one subblock and the same-different task (using previously unseen photographs). Within each subblock there were 20 trials at the family level and, for the Smiths, 36 trials at the individual level. Across sets of four subblocks, particular photographs were re-assigned randomly into the various conditions. Results will be presented in terms of "Blocks" of practice (equal to two 56-trial subblocks), which are approximately equivalent in number of individual level training trials with the Smiths (72) to a Session of Experiment 1a (88).

During training, an eyebrow mask was added to all images. This was different in shape from the Experiment 1 "headbands", and extended beyond the boundaries of the face region (see Figure 3.1e). To avoid subjects learning, for example, that the distance between the eyelids and the bottom of the mask was greater for Liz than for Ruth, the position of the mask was shifted up or down randomly within a 3 pixel range from trial to trial.

3.2.2 Experiment 2 - Results and Discussion

Table 3.6 shows mean identification accuracy during training. Learning of the individual Smiths was somewhat poor even in the upright orientation (cf. Table 3.1 for

Experiment 1)[5]. Inverted faces were learnt more slowly than upright faces. At least some learning in the inverted orientation did occur.

Table 3.6. Experiment 2: Average identification accuracy for the Smiths over the course of training. Each Block contains approximately the same number of individual level questions as a Session of Experiment 1.

	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8
Family								
Upright	83	96	97	96	95	97	99	98
Inverted	80	95	96	96	97	97	99	99
Individual (Smiths)								
Upright	59	61	77	76	79	78	81	80
Inverted	56	63	65	67	68	66	69	73

Note : Chance = 50%.

In assessing same-different performance only subjects who attained a criterion of 80% accuracy by the end of training (N=5 upright, and N=4 inverted) were included. In the upright orientation (Figure 3.5a), subjects performed the same-different task noticeably better than controls, as expected.

In the inverted orientation (Figure 3.5b), subjects showed two patterns of response. Three subjects generalised poorly to the same-different task; this argues that their good performance in the training phase had relied merely on learning of particular images. One person (Subject 3), however, generalised extremely well. This might, perhaps, be taken as evidence that covering the eyebrows during training had forced at least one person to learn a holistic representation of inverted faces. However this idea was refuted by her strategy report, which indicated a clear reliance on featural information. Specifically, in the absence of the eyebrows, Subject 3 had been able to find a tiny difference between Liz and Ruth's eyelashes. She reported that: "There was a notch in Ruth's towards the outer side of her face, her eyelashes were also much denser; Liz had a few clumps of eyelashes in the middle, her eyelashes were also much less

[5] There are a number of possible reasons for this. The eyebrow masks disrupt some holistic information by covering the connection between nose and brow-ridges, which might affect even upright subjects. Practice was also less spaced in time in Experiment 2 than in Experiment 1a and this may have produced poorer learning (cf. Donovan & Radosevich, 1999). It is also possible that subjects who knew they would be paid (in cash or course credit) regardless of how they performed, were less motivated than the personal friends of the experimenter tested in Experiment 1.

dense". Note that, as with eyebrow plucking, this "eyelash" strategy would hardly provide a viable method for identifying people outside an experimental setting.

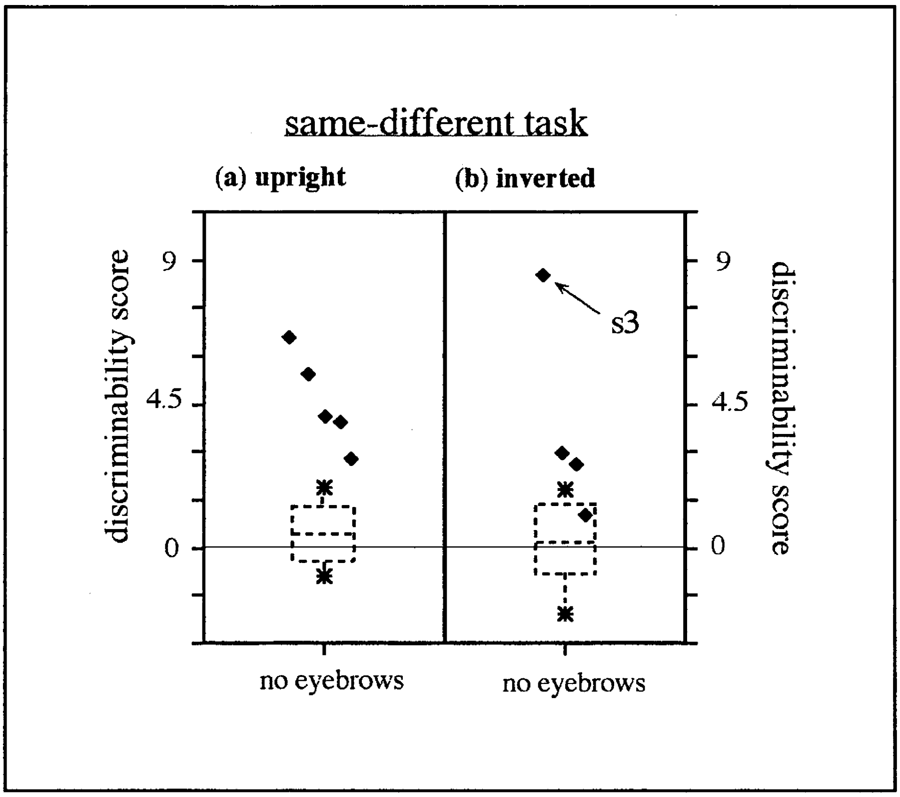


Figure 3.5. Experiment 2: Individual subject results in the same-different task for those who achieved the 80% accuracy criterion by the end of training. Control scores (taken from Experiment 1d) are shown as box plots with mean, ± 1 SD and range. One subject of particular interest is numbered (S3).

In summary, Experiment 2 tested learning when the easiest to find featural difference between the twins had been removed. There was no evidence, however, that this encouraged subjects in the inverted condition to form a holistic representation.

3.3 Experiment 3 - Composite test for holistic processing

So far, the only direct test of holistic processing for upright faces has been via the effects of removing one specific feature (the eyebrows). In Experiment 3, I used the composite paradigm (Young et al., 1987) to further assess whether trained twin discrimination had relied on holistic representations in Experiments 1 and 2. The composite paradigm is a well established direct test for holistic/configural processing. Subjects name a half-face of a familiar person, presented simultaneously with the other half of someone else's face; the two halves are presented either aligned or unaligned, as shown for the twin stimuli in Figure 3.6. The standard composite effect is that when faces are presented upright, half-faces (e.g., top half) are named slower in the aligned condition than in the unaligned condition; this interference from the irrelevant half indicates perceptual integration of the whole in the aligned stimulus. When faces are presented inverted, no composite effect occurs, indicating that the two halves are processed independently even when they are aligned.

In Experiment 3, I tested as many of the Experiment 1 and 2 subjects as possible, with each participating in the same orientation condition as before. My primary prediction was that the composite effect should be obtained in the upright orientation (indicating holistic processing) but not in the inverted orientation (indicating part-based processing). Possible overall differences in naming accuracy between top-half-faces and bottom-half-faces were also of interest, particularly for the inverted orientation, where the local features subjects had reported using were primarily in the top half of the face (i.e., eyebrows and eyelashes).

3.3.1 Experiment 3 - Method

3.3.1.1 Subjects.

Subjects were paid $10 for their participation. The delay since completion of their previous experiment was approximately 1 year for Experiment 1 subjects (upright and inverted), 3 months for the upright condition of Experiment 2, and 7 months for the inverted condition of Experiment 2.

There were several constraints on the number of subjects available for the composite task. Only Experiment 1 subjects could be tested on the Brown twins. From Experiment 2, I approached only subjects who had achieved 80% accuracy by the end of training. Finally, within Experiment 3 I used a brief retraining procedure to ensure that subjects could name whole-face stimuli accurately before proceeding to naming half-faces; here, I accepted only subjects who were at least 80% accurate for each twin by the end of two short re-training blocks. These various constraints left me with a total of N=6 for upright Browns, N=6 for inverted Browns, N=10 for upright Smiths, and N=7 for inverted Smiths.

3.3.1.2 Design.

Each subject participated only in their trained orientation condition. Face halves were referred to as "forehead" and "chin", to avoid confusion with "top" and "bottom" when faces were presented inverted. For a given block of composite trials, subjects named one half of the face for one set of twins (e.g., the chin half of the Smiths); this avoided confusion about what was to be named, and also allowed for binary key responses (e.g., Liz or Ruth). Order of half-to-name, and family for the Experiment 1 participants, was counterbalanced across subjects.

Within each block, aligned and unaligned trials were presented in random order. The comparison of interest came from aligned vs. unaligned composites of different twins (i.e., half of the composite from each sister). I also included same-twin composites (i.e., halves from two different photographs of the same person), for which data are not presented, to preclude a situation in which subjects could infer the identity of one half indirectly by knowing the other half (e.g., recognising the top half as Liz and realising that in all cases the bottom half had to be Ruth).

3.3.1.3 Stimuli.

The stimuli for the composite task were created from 12 previously unseen full-face (0°) photographs of each twin. These were treated as described in the General Method section of Experiment 1 and then cut in half horizontally. There were no eyebrow masks. In forming composites, each half face could be recombined with every other face from the same family, except for the one to which it originally belonged (e.g., Liz_1 top half could be combined with all Smith bottom halves except Liz_1). Note that there was no touching up of composite faces; the two halves were simply placed in the appropriate position relative to each other. In Young et al.'s (1987) stimuli, this procedure left an obvious join across the middle of aligned composites, making it clear to the subject where each half-to-name ended. The twin's faces were so similar to each other, however, that aligned composites connected almost perfectly. Thus, a marker line at the mid-point was added (see Figure 3.6).

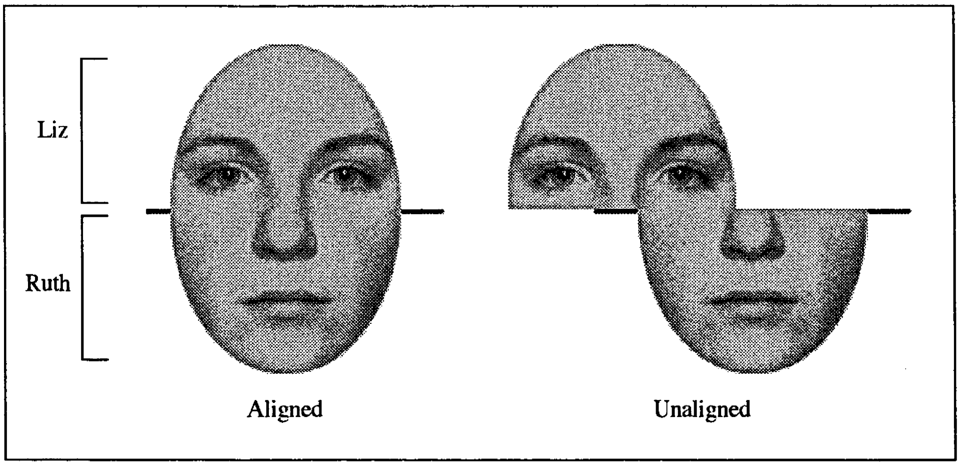


Figure 3.6. Examples of the composite faces used in Experiment 3.

Each condition of the composite test (e.g., naming "forehead" halves of Smith twins) included 128 trials. Of these, 64 comprised different-twin halves, and 64 comprised same-twin halves. Within each of these categories, 32 trials showed aligned stimuli, and 32 showed unaligned. For unaligned stimuli, the direction of offset was then varied (i.e., chin half moved left in 16 cases, and right in 16 cases). For each subject, the photographs appearing in each condition were randomly selected from the pool of 132 same-twin composites and 144 different-twin composites.

3.3.1.4 Procedure.

Experiment 1 subjects participated for 1.5 hrs (Smiths and Browns), and Experiment 2 subjects for 45 mins (Smiths only). Equipment was as in Experiment 1a, with responses recorded using a NewMicros button box measuring reaction times (RTs) accurate to 1 ms. The size of an aligned composite was equal to a whole face presented in Experiment 1a.

The retraining procedure was similar to the original training in Experiments 1 and 2, with two exceptions: only individual level categorisation was required; and the name verification procedure was replaced with a simple two alternative forced choice. On each trial, a stimulus face was presented until the subject responded (i.e., for the Smith twins, one button for Liz and another for Ruth), followed by a feedback stimulus (face plus correct name). Each short retraining block included 80 trials (e.g., 40 of Liz and 40 of Ruth), with photographs taken from the subject's original training set.

In the composite test, practice trials were given using male faces. Each trial in a given block then began with the word "chin" or "forehead" for 500 ms. The composite twin face (either aligned or unaligned) was then presented until the subject responded or for a maximum of 3,000 ms, followed by a 1,500 ms interval between trials. No feedback was given.

3.3.2 Experiment 3 - Results and Discussion

In the standard version of Young et al.'s (1987) paradigm, using famous faces, accuracy is high and thus the composite effect is assessed via reaction times. With the twin stimuli, however, I suspected that accuracy might be the more relevant measure. With accuracy, a composite effect takes the opposite direction from the usual effect with reaction times; that is, interference from the irrelevant half-face would be revealed as lower accuracy scores in the aligned condition than in the unaligned condition, rather than (or possibly in addition to) higher reaction time scores in the aligned condition.

Figure 3.7 shows accuracy for identifying half-faces in aligned and unaligned composites, separately for each set of twins (Browns and Smiths), and for "chin" and "forehead" halves of the face. Reaction times for correct responses are given in Table 3.7. Results are shown averaged over subjects.

Table 3.7. Experiment 3: Average reaction time (ms) for the composite task in each condition.

	Brown forehead		Brown chin		Smith forehead		Smith chin	
	aligned	unalign	aligned	unalign	aligned	unalign	aligned	unalign
Upright	874	847	1411	1282	1096	1109	1389	1315
Inverted	833	904	1164	1078	833	863	1444	1333

For the upright orientation (Figure 3.7a), a clear composite effect is apparent in three of the four conditions, with aligned less accurate than unaligned for naming the "chin" half of the Browns, $t(5)=2.61$, $p<.05$, the "forehead" half of the Smiths, $t(9)=1.99$, $p<.05$, and the "chin" half of the Smiths, $t(9)=1.99$, $p<.05$. The only exception was for naming the "forehead" half of the Brown twins, where accuracy in both conditions was at ceiling and so there was no room for any composite effect to emerge (reaction times were also very fast; see Table 3.7). In terms of individual subjects, the pattern was more variable (given the smaller number of trials) but six out of ten subjects showed a significant composite effect for at least one of the conditions tested ($\chi^2_{crit}(1, N=64) = 3.84$, $p < .05$). Also note that the composite effect did not reflect a speed-accuracy tradeoff. Averaged across all conditions, reaction times were 55 ms slower for aligned stimuli than for unaligned stimuli, and this composite effect on reaction times in fact reached significance for one condition (naming the "chin" half of the Browns, $t(5)=2.95$, $p<.05$). Thus, as predicted, the composite test demonstrated holistic processing for upright twin faces.

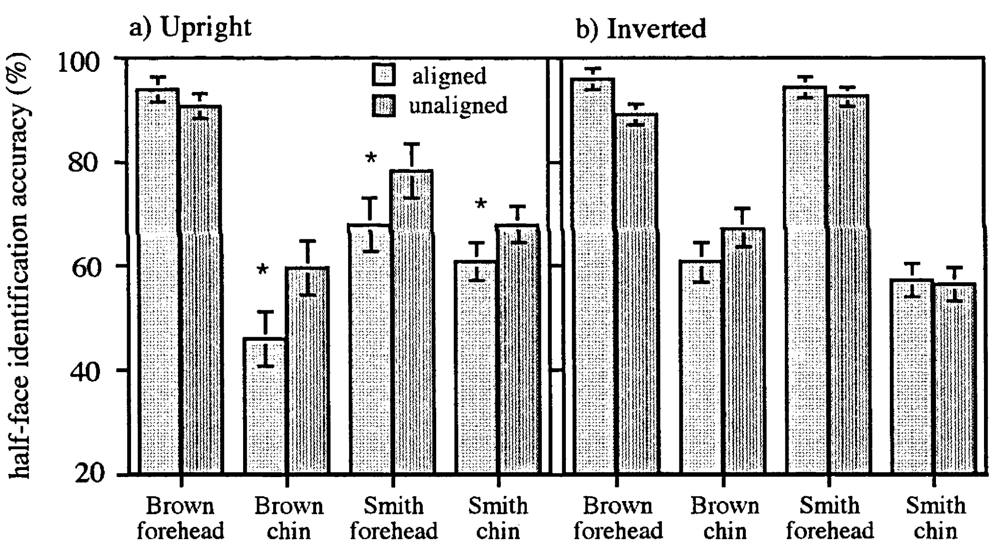


Figure 3.7. Experiment 3: Naming accuracy on the composite task for subjects in the upright (a) and inverted (b) conditions. Error bars are ± (√MSE / √N), appropriate for the within-subjects comparison of aligned versus unaligned; * = $p < .05$, one-tailed t-test.

In the inverted orientation, in contrast, no composite effect emerged. Figure 3.7b shows that three conditions produced effects trending in the wrong direction, and the one trend in the correct direction (the "chin" halves of the Browns) was not significant, $t(5)=1.78$, $p>.05$. In terms of individual subjects, there were no cases of a significant composite effect on accuracy for any subject for any condition ($\chi^2_{crit}(1, N{=}64) = 3.84$, all $p > .05$). Table 3.7 also shows no evidence for a composite effect on reaction times: aligned was only 23 ms slower overall than unaligned, and two out of the four conditions showed trends in the wrong direction.

For inverted faces, the difference in overall accuracy for the two halves of the Smith twins is also of interest. In the upright condition (Figure 3.7a), overall accuracy did not differ greatly between the "chin" half of the Smiths and the "forehead" half, but in the inverted orientation (Figure 3.7b), "forehead" halves of Smiths were identified at ceiling levels, while "chin" halves were close to chance. This pattern in the inverted orientation is consistent with previous evidence that the 7 individuals included in this test were all using information in the top half of the face (eyebrows in six cases, eyelashes in one case) to support discrimination of the Smiths. Interestingly, this includes Subject 12 from Experiment 1 who, in that experiment, failed to learn the Smith twins in 8 hours of training. During the re-training phase of the present experiment, however, his performance suddenly jumped from chance levels in the first retraining block to 90% correct in the second retraining block. According to his self-report, the reason for this was that, subsequent to Experiment 1, friends had informed him that the difference between the Smiths lay in their eyebrows, and he located the difference at the end of the first retraining block.

In summary, the composite test results provide direct confirmation that holistic processing occurred for upright twin faces (i.e., a composite effect), but not for inverted twin faces (i.e., no composite effect). It also again supported the conclusion from previous experiments, that subjects in the inverted condition were using featural information in the top half of the face.

3.4 General Discussion

The aim of the present study was to assess whether inverted faces could come to be processed holistically with practice. The current results argue strongly that this is not possible within the constraints of a training study. In terms of the amount of practice, I used a number of individually named trials with each twin which was easily greater than that used by Gauthier and Tarr (1997) in investigating face-like processing for greebles, and far greater than that shown by Tarr and Pinker (1989) to produce template-like representations of objects in new trained orientations. In terms of the style of practice, I presented each twin in multiple different photographs and across multiple different poses, thus giving subjects every opportunity to develop a generalisable whole-face representation. Finally, in using identical twins, I made learning based on local features as difficult as possible. This was very important. Had there been any obvious feature that discriminated the individual faces (e.g., light eyes vs. dark eyes), subjects in the inverted condition could have been expected to learn identity very easily (as occurred for the family decisions). In discriminating between identical twins, however, I provided maximum need for the perceptual system to develop a holistic representation of the whole face, if and where this were possible.

For inverted faces, no holistic processing emerged despite the careful selection of stimuli and practice level. There was no composite effect for inverted faces, directly confirming that processing remained part-based following practice. Instead, two basic patterns emerged. If a subject trained on inverted faces was able to find some tiny featural difference between the twins (usually after some effort), then the subject could (a) identify each twin reliably, and (b) generalise this ability to new photographs and new tasks, but (c) was unable to tell the twins apart if the preferred feature was not available in the stimulus (e.g., with the eyebrows covered in Experiments 1c and 1d, or with the wrong half-face in Experiment 3). The second pattern for subjects trained on inverted faces occurred when a subject was not able to locate a featural difference between the twins: here, the subject either failed to learn at all, or showed only non-generalisable learning of particular training-phase photographs.

The extraordinary reliance of inverted face learning on local features in the present experiments is perhaps most clearly emphasised by three individual cases. In Experiment 1, Subject 1 discovered the eyebrow difference in the Smith twins five hours into training, and suddenly jumped from chance levels to > 70% correct in a single session. Subject 12 from Experiment 1 failed to notice the eyebrow difference at all and correspondingly failed to reliably learn the Smiths for the entire eight sessions of training; in re-training for Experiment 3, however, he suddenly attained 90% accuracy after advice from other study participants on which feature to look for. In Experiment 2, with eyebrows not available, Subject 3 was the only person to find another featural difference between the Smith twins (eyelash clumping); correspondingly, she was the only subject in the inverted condition to show learning that could generalise well beyond the training phase photographs.

In contrast to this constant emphasis on tiny features for the inverted faces, strategy reports showed that most subjects who learned the twins upright did not even notice that these featural differences existed. Instead, upright subjects showed learning that generalised well and demonstrated two specific hallmarks of holistic processing: first, identification accuracy was unaffected by removal of a single feature from the face (eyebrows in Experiment 1; cf. similar results in Moscovitch et al., 1997); and second, half-face identification accuracy showed the standard composite effect (Young et al., 1987).

3.6.1 Possible caveats

I believe that I have made a compelling case against learning of holistic processing for inverted faces under the most conducive circumstances possible in a training experiment. However, I cannot, of course, conclude that there are no circumstances that could lead to development of holistic processing for inverted faces. The current training procedure, while using multiple poses and pictures, was still a rather limited version of real-world learning, in which thousands of different faces are seen in different poses, expressions, lighting conditions and so on, over a period of many years. It remains possible that a similar type of exposure to inverted faces could eventually lead to the development of holistic processing. I suggest that dentists and anaesthetists - who have at least moderate real-world exposure to inverted faces – might make interesting subjects for future investigation of this issue.

A more general caveat is that an inability to learn holistic processing in a novel orientation might perhaps arise from competition with any pre-existing holistic representation in the upright orientation, rather than being specific to faces. Thus, for example, if dog experts were able to develop a holistic representation of their breed-of-expertise in the upright orientation, it might be impossible for them to learn a subsequent holistic representation in the inverted orientation. To my knowledge this has not yet been tested.

3.6.2 Theoretical implications of the present results.

The primary importance of the present results is in showing that the "special" type of processing used for upright faces is not easily learned in the inverted orientation. This is an important finding for faces because of the apparent difference from the situation with other objects (at least in non-experts). For other objects, orientation effects disappear very quickly with practice (e.g., McKone & Grenfell, 1999), even when within-class discrimination is required (Tarr & Pinker, 1989; see also Tarr, 1995). This argues that it is easy to learn upright-like processing of nonface objects in new non-canonical orientations. The current results, however, indicate that the processing of faces remains highly orientation-specific even following substantial practice with the inverted orientation.

My results have less to say about the origin of "special" processing for faces. Had I shown learning of holistic processing for inverted faces in only 8 hours of practice (cf. the claims of the greeble studies), this result would have provided compelling evidence for the expertise hypothesis. The fact that I did not obtain this result does not allow the expertise hypothesis to be rejected, however, although note that my results are at least consistent with the alternative view that generic learning during adulthood is fundamentally different from learning of faces in infancy (de Hann, Humphreys & Johnson, 2002; Le Grand et al, 2001, 2003, 2004).

CHAPTER 4: WHICH TASKS BEST MEASURE "SPECIAL" PROCESSING FOR FACES?

4.1 Overview.

When choosing a task to evaluate whether face-like processing can sometimes emerge for objects (e.g., for human bodies or for objects-of-expertise), it is advantageous to know that under ordinary circumstances: (a) face-specificity has been fully demonstrated via comparison against multiple object classes with various different properties; and (b) the task produces a signature effect for faces that is completely absent for objects, rather than merely disproportionately large for faces. A review shows that the disproportionate inversion and part-whole effects satisfy (a) but not (b). My experiments then evaluate composite effects (a test of holistic/configural processing), plus contrast reversal. Both were tested for the first time with a natural object class (dogs). These and the tested inversion effect demonstrated face-specificity, but only the composite effect was completely absent for dogs. Note that, in the experiments described in the present chapter, all subjects were non-experts with dogs.

4.2 Introduction.

A long-standing idea is that faces are "special" in that upright faces are processed in a holistic/configural manner, while inverted faces, scrambled faces and objects are processed in a part-based manner. There is controversy regarding the exact nature of the processing used for upright faces (e.g., see Maurer, Le Grand, & Mondloch, 2002). Here, I conceptualise it as strong perceptual integration across the entire region of the face (excluding hair), that either involves no decomposition into parts beyond simple image components processed in early vision (a more "holistic" interpretation; Farah, 1996; Moscovitch, Winocur, & Behrmann, 1997; Tanaka & Sengco, 1997), or else interactions between multiple parts across the face (a more "configural" interpretation; Rhodes, 1988). I do not distinguish between these two slightly different interpretations, and use the term holistic throughout as shorthand for "the holistic/configural processing style apparently special to upright faces".

In this chapter, I examine several tasks that have been associated with face-specific processing. My aims are (a) to determine the extent to which the phenomena revealed in these tasks are truly specific to faces, and (b) to evaluate the extent to which the phenomena are theoretically associated with holistic processing. This work is important because of the use, or potential use, of these tasks in evaluating theories of the origin of "special" processing for faces. While it might be that faces are special due to some innate representation of structure (Morton & Johnson, 1991) and/or preferential exposure to faces during a critical period in early infancy (Le Grand, Mondloch, Maurer, & Brent, 2001, 2004), the alternative theory is that generic expertise in individual-level identification (eg. dog 1 vs. dog 2, by analogy with Bill vs. Sam for faces) leads to the development of special processing for faces. The method of evaluating this expertise hypothesis (Diamond & Carey, 1986; Gauthier & Tarr, 2002; Meadows, 1974) then relies on testing experts in non-face object domains (e.g., dog-show judges looking at dogs) for the phenomena suggested to be associated with face-specific processing in the paradigms discussed here. It is therefore essential to know that these phenomena produce a clear dissociation between faces and objects in novices: that is, subjects who have some level of general familiarity with the object class tested, but are not expert at making within-class discriminations (e.g., have poorer discrimination of two dogs than of two faces).

4.2.1 <u>Disproportionate inversion effect on memory.</u>

Most objects are somewhat harder to remember when learned and tested inverted than when learned and tested upright. However, across many studies it has consistently been found that the inversion effect on recognition memory for faces (commonly 20%-25% decrement) is much larger than for objects (usually 0%-8% decrement). This disproportionate inversion effect for faces has been reported in comparison to aeroplanes, period costumes, houses, stick figures, buildings, landscapes, dog faces, and side-on views of dogs of several breeds (Diamond & Carey, 1986; Scapinello & Yarmey, 1970; Yin, 1969). The large inversion effect is usually interpreted in terms of holistic processing assisting memory for upright faces, but not inverted faces or objects in either orientation.

As a method of assessing face-specific processing, the strength of the inversion paradigm lies in the wide range of other objects to which faces have been compared. It is difficult to choose any one object class as the perfect control stimulus for faces. It is generally accepted that all members of the object class should share a common basic structure (a "first order" configuration, eg. all dogs have a head attached to a body above four legs; Diamond & Carey, 1986). Beyond this, however, there is no single object class that exactly matches faces in terms of complexity, spatial frequency components, degree and type of variability between individuals, surface texture, clarity of part boundaries, natural versus manmade status, and so on. Thus, the fact that faces produce large inversion effects in comparison to a broad range of objects is very important in establishing face-specificity.

The inversion effect on memory unfortunately also has two significant disadvantages as a measure of face-like processing. The first is that the signature of face-specific processing is a disproportionate inversion effect, rather than the presence as opposed to absence of an inversion effect. This means that, when testing experts for example, the question tested must be whether objects-of-expertise produce an inversion effect as big as for faces, or substantially larger than in novices, rather than whether they produce an inversion effect at all.

The second disadvantage is that there is no direct theoretical logic linking disproportionate inversion effects with holistic processing (Valentine, 1988, 1991). Yin (1969) was correct in his suggestion that upright faces are processed holistically, while inverted faces are not, but the evidence for this came from other paradigms discussed below. Thus, for objects, even if a face-sized inversion effect on memory were found, this would not, in itself, be sufficient to demonstrate holistic processing.

4.2.2 The part-whole paradigm.

The part-whole effect (Davidoff & Donnelly, 1990; Tanaka & Farah, 1993) has a logical connection to holistic processing in that it examines the effect of surrounding context from the whole on memory for a part. In Tanaka & Farah (1993; also see Tanaka & Sengco, 1997), subjects first learned whole faces (e.g., John). In a subsequent memory test, stimulus pairs were presented, either as isolated parts (e.g., John's nose vs. Bill's nose) with subjects asked to choose the studied part (John's nose), or in the context of the whole face (e.g., John's nose in John's face vs. Bill's nose in John's face) with subjects asked to choose the studied face (John's face). In the upright orientation, memory for the face part was substantially better in the whole face condition than in the isolated part condition, arguing for some form of holistic processing.

Several studies (Davidoff & Donnelly, 1990; Donnelly & Davidoff, 1999; Tanaka & Farah, 1993) have confirmed that inverted faces and scrambled faces do not produce any part-whole effect. This same finding – namely a zero part-whole effect – has been shown for houses in the specific methodology used by Tanaka and Farah (1993; Donnelly & Davidoff, 1999). However, although results of this one object class suggested that the part-whole effect might provide a presence versus absence difference from faces, tests with other objects did not replicate the finding. Car fronts, biological cells, and dog faces (Tanaka et al. 1996, cited in Tanaka & Gauthier, 1997), and the artificial objects "greebles" (Gauthier & Tarr, 1997) have all demonstrated small part-whole effects. Even houses have produced a part-whole effect, using an immediate-memory version of the paradigm that avoids repetition of particular parts across more than one study exemplar (Donnelly & Davidoff, 1999). The same result has also been obtained for chairs (Davidoff & Donnelly, 1990). Across studies it is clear that objects show a nonzero part-whole effect, but one that is substantially smaller than that for faces; that is, there is again a merely a disproportionately large effect for faces.

Why might a method apparently tapping holistic processing still produce some effect for objects? There are three *a priori* possibilities. The first is that a small amount of "face-specific" holistic processing occurs for objects. Results of other paradigms, however, argue against this idea (see below and the General Discussion). Second, it could be that the part-whole effect does not actually tap face-like holistic processing, but instead taps more general context, gestalt, or global processing effects. As a methodology, the part-whole effect seems related to word superiority and shape superiority paradigms (Davidoff & Donnelly, 1990). In the word superiority effect, a degraded letter is identified more accurately in a word (e.g., the third letter in knit vs. knot) than alone (i vs. o) or in a random letter string (gjik vs. gjok; Reicher, 1969; Wheeler, 1970). Similarly, the angle of a line is more accurately perceived in the context of a connected shape (e.g., a parallelogram) than when tested alone (e.g., Enns & Gilani, 1988). The advantage of whole over parts for objects might then reflect similar generic context effects, in which processing of any type of part is enhanced by placing it in a coherent or familiar structure. Under this view, the part-whole effect for faces would arise from the same generic source. This idea, however, fails to explain why the part-whole effect is much larger for faces than for objects. A third alternative is that the part-whole effect measures contributions from two sources: it arises partly from face-specific holistic processing, and partly from generic context or global processing effects. This view can explain both the fact that objects show a nonzero part-whole effect, and the fact that it is much larger for faces.

Even under this last view, the part-whole effect does not provide a pure measure of face-specific holistic processing. Instead, as with the inversion effect, any claim of face-like processing for objects (e.g., for objects-of-expertise), would always require demonstrating a part-whole effect that was as large as that for faces, or at least much larger in experts than novices.

4.2.3 The composite effect.

In Young, Hellawell, and Hay's (1987) composite paradigm using famous faces, the top half of one person's face was combined with the bottom half of a different person's face (e.g., Tony Blair's forehead with George Bush's chin). For upright faces, when these two halves were physically aligned, subjects were slower to name either half (e.g., the top half) than when the two halves were offset, indicating holistic processing when the halves were aligned. For inverted faces, in contrast, there was no difference in naming times for aligned and unaligned stimuli, showing no holistic processing (also see Carey & Diamond, 1994). This pattern has been confirmed for novel faces using a same-different judgement and aligned vs. unaligned composites (Le Grand et al., 2004). A composite effect for novel faces has also been demonstrated indirectly by Hole (1994), who tested only aligned composites and relied on the difference between upright and inverted conditions being in the opposite direction to usual (i.e., inverted was better than upright) to argue that holistic processing must have occurred for upright composites.

From a theoretical perspective, the composite paradigm in its aligned vs. unaligned version would seem to provide a particularly strong way of tapping perceptual fusing of parts into a new whole. This is because (a) unlike the part-whole effect, the context provided by the other half is present on the screen in both aligned and unaligned conditions, and (b) simple response competition from the two halves (one suggesting "Blair", one suggesting "Bush") is also matched across the two conditions. Despite a general appreciation of the theoretical value of this paradigm for faces (e.g., Maurer, et al., 2002), it has been tested on only one class of objects, namely greebles. Results showed no composite effect in greeble novices; indeed, the effect was approximately a 2% (and 42 ms) difference in the reverse direction for a composite effect (Gauthier & Tarr, 2002). This raises the possibility that, unlike inversion and the part-whole effect, the composite paradigm might produce an absence of its signature phenomenon for objects, rather than merely a disproportionately large effect for faces. To support this conclusion, however, it would be necessary to have data on other object classes.

4.2.4 Contrast reversal effects.

A final test that has been suggested to produce face-specific effects is the contrast reversal paradigm. A contrast reversed image is one in which the luminance values of pixels have been reversed so that the picture looks like a photographic negative. (This manipulation is also sometimes referred to as contrast inversion, but "reversal" is used here to more clearly distinguish it from the change in orientation meant by the term "inverted face".)

Theoretically, I do not see contrast reversal effects as arising from holistic processing for faces. Instead, in agreement with other authors, I presume that contrast reversal for faces disrupts information about shape-from-shading extracted in mid-level vision, rather than tapping holistic processing in high-level vision. In support of this interpretation, Kemp, McManus, and Pigott (1990) showed that inversion and contrast reversal each reduced accuracy in a match-to-sample task, but did so in an additive fashion (i.e., inverted faces were as strongly affected by contrast reversal as upright faces), arguing that inversion and contrast reversal effects arise from different stages of the visual processing stream (also see Bruce & Langton, 1994). Further, Hole, George & Dunsmore (1999) found a composite effect (i.e., evidence of holistic processing) for contrast reversed faces showing that the two effects are independent.

Although there is no link between contrast reversal effects and holistic processing, it is of interest that contrast reversal effects for faces are very large (e.g., Bruce & Langton, 1994; Johnston, Hill, & Carman, 1992; Kemp, Pike, White, & Musselman, 1996). This has led to the suggestion that sensitivity to contrast reversal might be a phenomenon specific to faces in novices, and so worth evaluating in experts (Gauthier, Williams, Tarr & Tanaka, 1998). I have some doubts that contrast reversal will always be face-specific even in novices (the idea that recognition of complex scenes is sensitive to contrast reversal will have occurred to anyone who has looked at photographic negatives of their holiday snap-shots), but empirical evidence with objects tested to date is consistent with face-specificity. In the only published study, Gauthier et al. (1998) found no effect of contrast reversal for identifying greebles. In two unpublished conference presentations, Subramaniam and Biederman (1997) reported the same result for chairs, and Nederhouser, Mangini, Biederman, and Kazunori (2002) the same result for "blobs" (pictures of smooth three-dimensional objects with a number of random looking protrusions).

The size of the contrast reversal effect for objects in these prior studies has been essentially zero, rather than merely smaller than for faces. This potentially indicates an important all-or-none signature of face-specific processing. However, the zero contrast effect might be partly attributable to the type of objects tested. Chairs and greebles are both artificial objects with relatively clear part-boundaries that could reduce reliance on shape-from-shading information. Although 'blobs' were introduced in an attempt to deal with this criticism, the blobs had little in the way of surface texture (unlike faces).

4.2.5 The present study.

Of the techniques commonly used as measures of holistic processing in faces, two – the inversion effect on memory and the part-whole task – have been clearly established to produce face-specific effects. For both these tasks, faces have been contrasted with a broad range of objects, differing in many properties. Unfortunately, however, both inversion effects on memory and the part-whole effect show only a quantitative rather than a qualitative difference between faces and objects. That is, the effects are disproportionately large for faces, but are not absent for objects. This means that the simple presence of one of these effects cannot be used to establish holistic processing (e.g., in objects-of-expertise).

The other effect that appears to be associated with holistic processing – the composite effect –has strong theoretical logic linking it to perceptual integration across the whole of the stimulus[1]. Its empirical status as producing face-specific effects,

[1] One further paradigm that has been associated with holistic processing in faces is relational versus local alteration, where faces are varied either on the spacing between face parts (e.g., nose-mouth distance), or on local part information (e.g., eye colour, lower lip shape). Small spacing changes are detected much more poorly in inverted faces than in upright faces, but local changes are less sensitive or insensitive to inversion (Gilchrist & McKone, 2003; Leder & Bruce, 1998; Le Grand et al., 2001). The "second-order" relational processing for upright faces demonstrated in these paradigms is considered to be a key part of face-specific holistic/configural processing, but I have not discussed this paradigm here because it does not transfer neatly to objects. For example, it is not clear whether increasing the distance between the forelegs and hind legs in a dog should be treated theoretically as a relational change (altered spacing), or as a part change (longer torso).

however, has not been fully established; it has been tested only for greebles, and not for any natural objects.

The primary aim of my experimental work was thus to test the composite task for a class of natural objects (dogs). I was also interested in testing the effects of contrast reversal on a natural stimulus class. While there is no single "perfect" control stimulus for faces, dogs were chosen as well-matched to faces on a number of important variables. Both have a canonical upright, and individual members of the class all share a first-order configuration. Both are natural stimuli for which differences between individuals arise from genetic variability; this genetic component means that, unlike manmade objects such as greebles, individuals differ from each other on a great many dimensions at once (length of chin to shoulder, <u>and</u> length of hip to foot, <u>and</u> size of foot, <u>and</u> so on) rather than just three or four. Finally, both dogs and faces have some degree of surface texture, and include fuzzy part boundaries (eg., the neck merges smoothly into the torso for dogs, just as the nose merges smoothly into the cheek in faces). The dogs used were labradors. These have even colouring across their bodies, and were chosen to avoid introducing a deliberate emphasis on parts such as might occur for breeds with strong colour boundaries (e.g. beagles).

The core results are presented in Experiment 6 (composite task). Inversion effects on memory are also tested in Experiment 4, to ensure I could replicate the standard finding of a larger inversion effect for faces than objects for the current stimuli. In Experiment 5, I examined contrast reversal effects for dogs. Although there is no evidence that contrast reversal has anything to do with holistic processing, I thought it worthwhile assessing whether contrast reversal effects remain face-specific when compared to a natural stimulus with fuzzy part boundaries and some degree of surface texture (i.e, an object class in which shape-from-shading information is potentially important).

4.3 Experiment 4 – Inversion Effects On Recognition Memory

In recognition memory tasks, faces usually show larger inversion effects than other objects even when the task requires within-class processing (e.g., Yin, 1969). Experiment 4 aimed to replicate this finding for the current labrador stimuli.

An important factor affecting memory performance is the similarity between the items in a given class, with better memory if the stimuli are more different from each other than if they are more similar. To equate internal similarity between the face stimulus set and the dog stimulus set, I took the approach of matching performance levels in the inverted orientation. In this orientation, subjects should not be able to apply special processing mechanisms to either faces or dogs.

4.3.1 Experiment 4 - Method

4.3.1.1 Subjects.

Twenty-two Caucasian subjects (11 males) completed the memory task for course credit (N=14) or $3 (N=6) for the 15min experiment. Most also participated in other unrelated experiments in a longer session. Age ranged from 18-43, with most subjects aged 18 years (only two were over 35 years). All reported normal or corrected-to-normal vision.

4.3.1.2 Design.

Stimulus class (dogs vs. faces) and orientation (upright vs. inverted) were manipulated within subjects. For each stimulus class and orientation (e.g. upright dogs) there was a study (learning) phase, followed by a distracter phase, and then the memory test phase. Order of stimulus class and orientation conditions was counterbalanced across subjects. On each test trial, subjects saw two dogs (or two faces), one of which had appeared at study ("old") and one of which was unstudied ("new"), and were asked to select the old one. The dependent measure was accuracy on this two alternative forced choice (2AFC) recognition memory task (chance = 50%).

For the purposes of the 2AFC presentation, 60 dogs (and 60 faces) were organised into 30 pairs, with appearance and stance matched as closely as possible across the exemplars in each pair. Within each stimulus class, subjects received half of the pairs in the upright condition and the other half in the inverted condition with the particular set received in each orientation counterbalanced across subjects. Within each pair, one of the two items was designated as studied while the other remained unstudied. Assignment to studied/unstudied status was counterbalanced across subjects. Thus for each condition (e.g., upright dogs) 15 exemplars were studied with 15 corresponding pairs presented at test.

4.3.1.3 Stimuli.

Dog stimuli were photographs of 60 yellow labrador retrievers, shown standing in side-on pose (i.e. a standard show-dog pose). Examples are shown in Figure 4.1, and the full set is presented in Appendix I. There were 16 female labradors and 44 males (all are referred to as dogs). Dogs were taken from a mixture of sources including books (The Book of the Labrador Retriever, Nicholas, 1983; The Labrador Retriever Club of Victoria Inc's "Gold Book") and breeder web-sites in the public domain. Most of the dogs were American (approximately 38) or Australian (14) with the rest English or Canadian. There was a range of lighting direction in the photographs, as well as variability in image quality. Dog pictures were converted to greyscale and scaled to 4.9 cm – 6 cm from nose to tail (average 5.7 cm) by 3.5 cm - 4.6 cm from top of head to paws (average 4.2 cm). At the experimental viewing distance of approximately 45 cm, images subtended visual angles of 7.2° horizontal and 5.3° vertical on average. Any extraneous information (e.g., grass stalks over paws; handler's fingertips on tail) was edited out of the image. Half of the dogs faced to the left and half faced to the right.

Face stimuli were photographs of Caucasians. These were taken from a mixture of different online databases in the public domain (Stirling PICS: Nottingham-scans; University of Ljubljana CVL and CV, PTER, Velenje; Max-Plank Institute for Biological Cybernetics, Tuebingen, Germany), to match the variability in lighting direction and image quality in the dog stimulus set. To match the variability in sex, there were 10 females in the total of 60 faces. Example faces are shown in Figure 4.2, and the full set is presented in Appendix II. All pictures were front view with neutral expression, and contained no extraneous information such as glasses or beards. Each face was cropped by hand to exclude hair, but to retain as much forehead, cheek and chin shape as possible. This was necessary for the composite task (Experiment 6) and also meant that, as with the dogs, the outline shape varied from face to face. After cropping, each face image was sized to 3.1 cm - 3.8 cm at the widest point (average 3.4 cm) by 4 cm - 4.6 cm at the tallest point (average 4.4 cm), corresponding to an average of 4.3° by 5.6° at 45 cm viewing distance. Small blemishes on the faces were edited out.

All stimuli were placed on a neutral grey background. Brightness and contrast were equated within each stimulus class set as far as possible, but there was still quite a lot of variation. Inverted stimuli were created by rotating the pictures 180°. All manipulations were done using Adobe Photoshop (5.5) software.

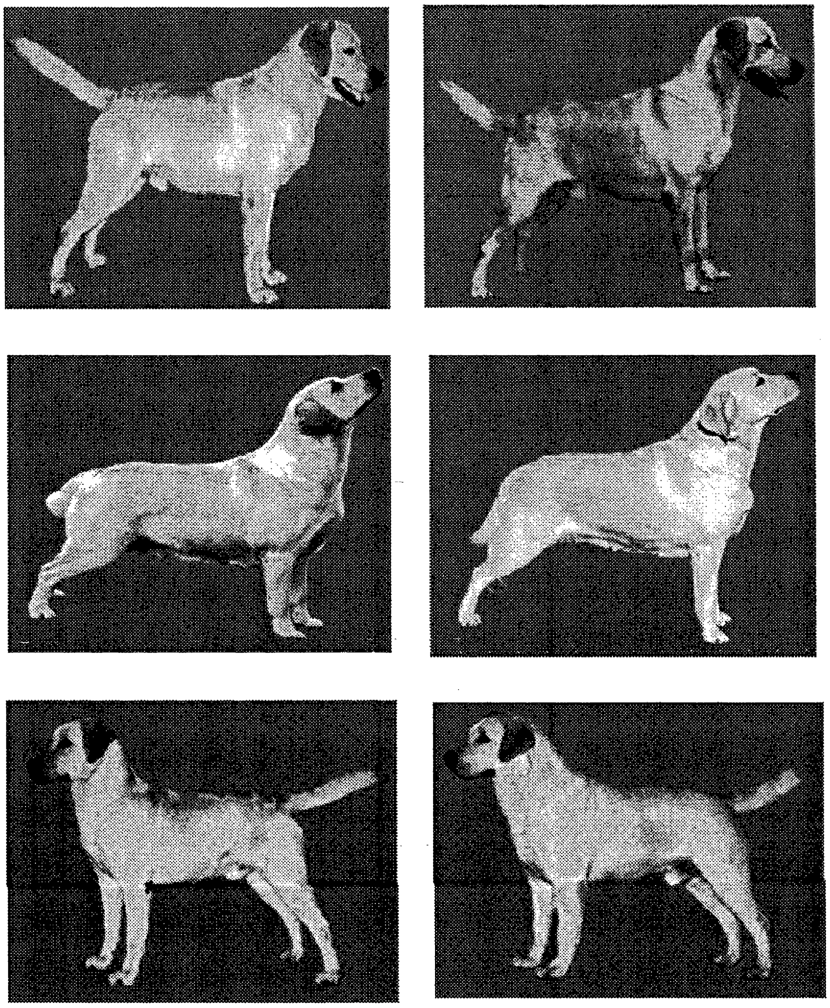


Figure 4.1. Examples of the range of dog stimuli used in Experiments 4-6, shown here paired as for the test phase in the memory task (Experiment 4).

It was impossible to obtain two different photographs of each of 60 labradors, and thus the same images were used at study and test for both dogs and faces. The dog recognition task could then technically be considered a picture recognition task (as is true of all studies reviewed in this article); however, note that the close pairing of dogs within test pairs, combined with the multiple slight differences in appearance across all 60 dogs, makes it very unlikely that accurate memory could rely on a single local cue (e.g., angle of tail) or purely image-based information (exact level of contrast). Also note that faces were similarly shown in the same image at study and test, meaning that any differences between dogs and faces could not be attributed to picture repetition.

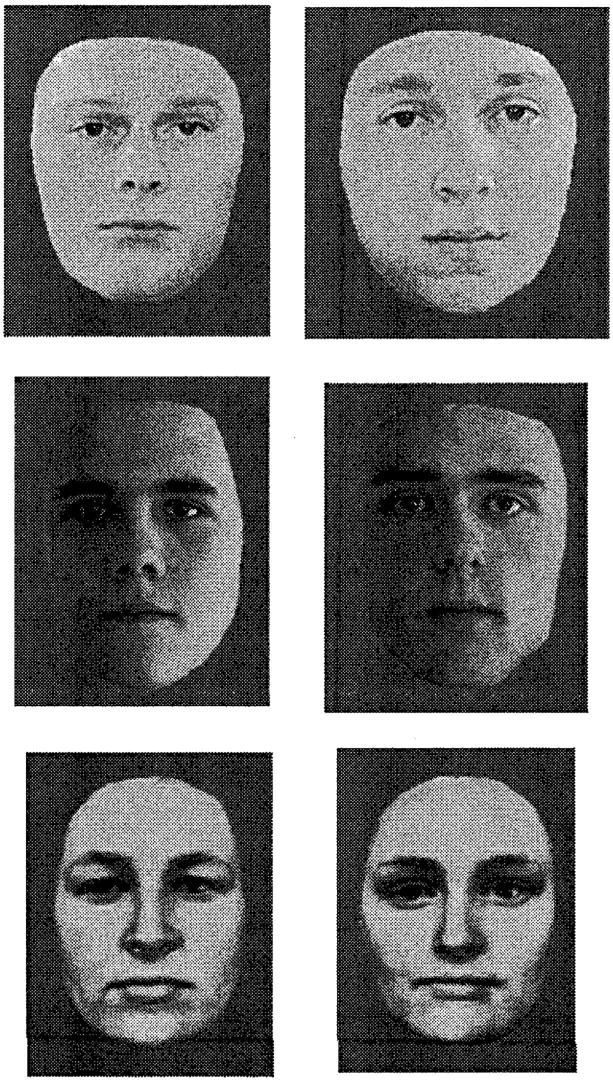


Figure 4.2. Examples of faces from each of the databases used in Experiments 4-6 (shown here paired as described in Figure 4.1).

4.3.1.4 Procedure.

Pictures were presented on an iMac computer, with a 17" monitor, using PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). Responses were recorded via the keyboard.

During each study phase, dogs or faces were presented one at a time in the centre of the screen. Subjects were instructed to learn these for a later memory test. Each picture was presented for 5000 ms, with an intertrial interval of approximately 575 ms. Presentation order was randomised for each subject.

During each distracter phase, subjects were presented with multiplication problems on screen and were instructed to answer, on paper, as many as they could in 1 min. No subject completed all the problems in this time.

In each memory test phase, the two stimuli presented on each trial were shown simultaneously at the same height and 13.3 cm (16.8°) apart centre to centre. Subjects pressed the "z" key to indicate that the left exemplar was "old" and the "m" key to indicate that the right exemplar was "old". The old exemplar appeared on the left and right equally often. Viewing time was unlimited and the intertrial interval was approximately 100 ms.

4.3.2 Experiment 4 - Results

The percentage of trials on which the old stimulus was correctly chosen was calculated for each subject for each condition. Figure 4.3a shows scores averaged across subjects for faces and dogs. A first point is that, as intended, performance for inverted faces and inverted dogs did not differ, $t < 1$, arguing that the two classes had equal levels of internal similarity. However, when the stimuli were upright, and subjects should be able to apply face-specific perceptual mechanisms, faces were remembered much better than dogs, $t(22)=3.25$, $p < .01$. A 2-way repeated-measures Analysis of Variance (ANOVA) confirmed a larger inversion effect for faces than for dogs, via a significant interaction between class (faces vs. dogs) and orientation (upright vs. inverted), $F(1, 21) = 5.74$, $MSE = 186.01$, $p < .05$. Memory was better for upright than inverted faces, $t(21) = 5.10$, $p < .001$, and the smaller inversion effect for dogs was also significant, $t(21) = 2.41$, $p < .05$.

A possible concern is that the smaller inversion effect for dogs than faces could reflect less room to show a decrement because of poorer upright performance in the former case. To anticipate this concern, two brief additional experiments were conducted[2]. These tested dogs only, and varied overall performance via changes in the memory set size. In the first (N=14), the 15 to-be-learned dogs in a given orientation were split into two study-test cycles of 8 dogs (i.e., the subject learned 8 dogs and was tested on 8 pairs) and then 7 dogs (the subject learned the remaining 7 dogs and was tested on 7 pairs). This produced better upright performance but still only a small inversion effect (Figure 4.3b). In the second (N=16) the 15 dogs were split into three study-test cycles of 5 dogs each. This improved performance for upright dogs (Figure

---

[2] These were designed by me, but subject recruitment and testing was conducted by a research assistant (Jacqui Brewer).

4.3c) to approximately match that for upright faces in the original experiment (Figure 4.3a), but again produced only a small inversion effect for dogs.

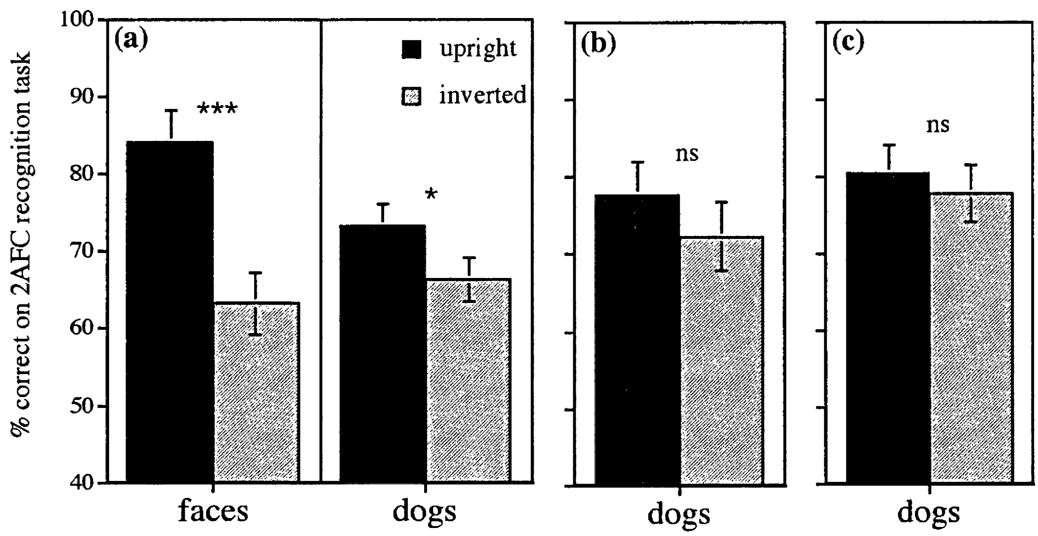


Figure 4.3. Experiment 4: Recognition memory test, (a) faces and dogs (set size = 15, n = 22), (b) dogs (set size = 8 + 7, n = 14) (c) dogs (set size = 5 + 5 + 5, n = 16). Error bars are the equivalent to SEM for making the within subjects comparison between upright and inverted orientations. *** $p < .001$, * $p < .05$, ns $= p > .05$.

4.3.3 Experiment 4 - Discussion

The current results confirm a disproportionately large inversion effect on memory for faces as compared to the labrador stimuli. The size of the inversion effect for dogs across the three tests (Figures 3a-3c) fell in the usual range of approximately 0%-8%. Thus this experiment replicates many previous studies contrasting faces with a broad range of manmade and natural objects.

4.4 Experiment 5 – Contrast Reversal

Face recognition is very adversely affected by contrast reversal, namely swapping luminance values in a picture so that it looks like a photographic negative (e.g., Galper, 1970). In Experiment 5 the contrast reversal effect for dogs was tested.

While contrast effects have previously been examined for greebles and (in unpublished conference presentations) for chairs and "blobs", the current labrador stimuli provided the first test of contrast reversal effects for a natural as opposed to manmade object class. The effects of orientation inversion and its interaction or otherwise with contrast reversal were also examined; this has not previously been done for any object class.

The task was simultaneous-presentation identity matching, in which the two stimuli could be either (a) both original contrast, (b) both reversed contrast or (c) one reversed and one original contrast. Stimulus presentation time was limited and accuracy was used as the measure. I predicted that faces would be affected by contrast reversal and also independently by inversion (Bruce & Langton, 1994; Kemp et al., 1990). The question of interest was whether dogs would show the same pattern.

4.4.1 Experiment 5 - Method

4.4.1.1 Subjects.

Twenty new Caucasian subjects (5 male) completed the experiment for course credit (N=5) or $5 for the half-hour experiment. Age ranged from 18-30 years, and all subjects reported normal or corrected-to-normal vision.

4.4.1.2 Design.

Class (dogs vs. faces), orientation (upright vs. inverted) and contrast condition (both original vs. one original one reversed vs. both reversed) were all manipulated within subjects. Class and orientation were blocked (e.g., a block of upright dogs) with block order counterbalanced across subjects. The order of contrast condition was randomised for each subject within each block. On each trial a pair of dogs (or faces) was presented simultaneously for a brief period. The dependent measure was accuracy to judge whether the exemplars were of the same identity or different identity, regardless of whether the contrast was the same or not.

4.4.1.3 Stimuli.

The stimuli were created from the 60 faces and 60 dogs used in Experiment 4. For each item a contrast reversed version was created using the Inverse function in Photoshop. Examples are shown in Figure 4.4. For same identity pairs both stimuli showed the same dog (face) exemplar, in one of the three contrast combinations. All 60 possible same identity pairs were created in the format where both stimuli were original contrast, and also in the format where both stimuli were reversed contrast. In the format where one of the stimuli was original and one contrast reversed the reversed item could appear on either the left or the right giving 120 possible same identity pairs. For each subject, 30 both original pairs, 30 both reversed pairs and 60 one original one reversed same identity pairs were chosen at random from the full set to be presented in each orientation condition. For different identity pairs, each trial showed two different but similar exemplars (the pairs developed for the memory task in Experiment 4 were used). To create the contrast reversal conditions the procedure used for the same identity pairs was repeated. In total there were 240 trials in a block.

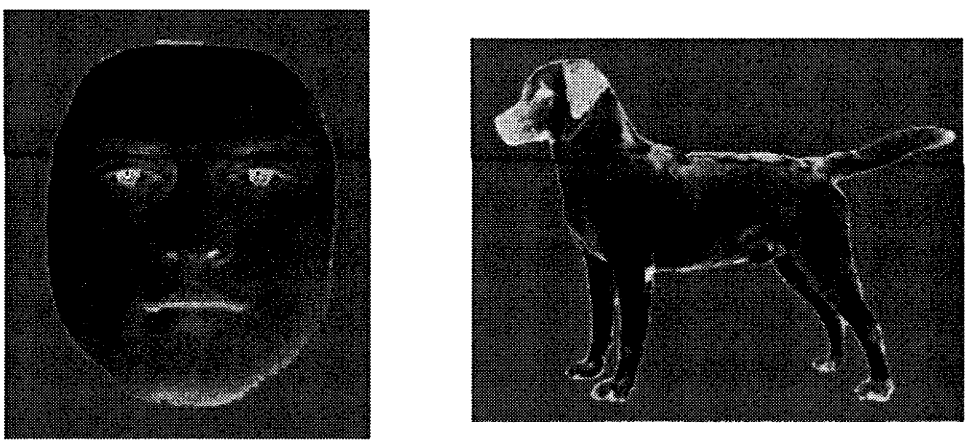


Figure 4.4. A contrast reversed face and dog as they appeared in Experiment 5.

4.4.1.4 Procedure.

Apparatus was as for Experiment 4, with the addition of a NewMicros Button Box connected to the iMac via a Keyspan USB twin serial adaptor. Subjects had the task explained to them, and were given 10 practice trials on a version of the task using chairs as stimuli (chair pictures courtesy of Bruno Rossion & Michael Tarr). On each trial, the stimulus pair was presented until the subject responded or for 600 ms. This brief presentation allowed subjects to look at each exemplar only 1-2 times because stimuli were far enough apart that eye movements were required to foveate each stimulus. Subjects were warned that they might feel they were guessing on some trials but were required to enter a response anyway. A green/right button indicated same identity, a red/left button indicated different identity. After response, there was a blank screen for at least 200-400 ms (longer if the subject responded before the stimulus had disappeared from the screen). Subjects were told that on 50% of the trials, the pair would have the same identity and that on 50% of the trials they would have different identities.

On each trial, the position of each stimulus on the screen was jittered. Jittering was used to ensure that subjects could not learn to spatially focus their attention on particular regions of the screen before the stimuli were presented, thus discouraging decisions based on a single local part of the stimuli (e.g. the dogs' tails). Each exemplar of a pair could appear in a one of five positions. For dogs these ranged from 7.5 cm to 8 cm (9.5° to 10.2°) apart horizontally from centre to centre and 2 mm to 5 mm (0.3° to 0.6°) apart vertically. For faces these ranged from 8 cm to 11 cm (10.2° to 13.9°) apart horizontally from centre to centre and 2 mm to 1 cm (0.3° to 1.3°) apart vertically. The actual position selected was chosen randomly for each stimulus on each trial, with the constraint that the two exemplars were never presented at exactly the same height.

4.4.2 Experiment 5 - Results

Mean accuracy in the same-different task for all conditions is presented in Table 4.1. The both original and both reversed conditions were focused on first. Comparing these conditions allows the best assessment of the extent to which contrast reversal damages stimulus processing, because the conditions are otherwise equivalent; specifically, for same-identity trials the stimuli in a pair are physically identical in both cases. (Comparing the both original condition with the one-original-one-reversed condition is not valid because, in the latter condition, stimuli in a pair are never physically identical). Results for these conditions collapsed across same and different trials are plotted in Figure 4.5 and analysed below.

A repeated measures ANOVA including class (dogs vs. faces), orientation (upright vs. inverted) and contrast condition (both original vs. both reversed) revealed no three-way interaction, $F(1, 19) = 1.66$, $MSE = 16.74$, $p > .2$, but class showed significant two-way interactions with both orientation, $F(1, 19) = 8.15$, $MSE = 41.74$, $p < .01$, and contrast condition $F(1, 19) = 39.97$, $MSE = 25.86$, $p < .001$. Thus, further analysis was conducted for faces and dogs separately.

Table 4.1. Experiment 5: Mean same-different accuracy (% correct and SEM) for dogs and faces in each contrast condition (collapsed across same and different trials).

	upright			inverted		
	Both original	One original one reversed	Both reversed	Both original	One original one reversed	Both reversed
Faces	85.17	63.42	68.83	76.00	59.79	63.25
	(1.58)	(1.46)	(2.19)	(1.89)	(1.42)	(1.97)
Dogs	74.33	66.04	69.83	72.67	65.04	68.42
	(2.36)	(1.39)	(1.837)	(1.76)	(1.84)	(1.67)

For faces, two results are apparent in Figure 4.5a. First, with respect to the effects of contrast reversal, same-different accuracy was substantially higher in the both original condition than in the both reversed condition. Second, with respect to orientation, performance was worse for inverted faces than for upright faces for each of the contrast conditions. A 2-way ANOVA confirmed a main effect of contrast, $F(1, 19) = 150.09$, $MSE = 28.18$, $p < .001$, and a main effect of orientation, $F(1, 19) = 41.89$, $MSE = 25.97$, $p < .001$. There was no interaction, $F(1, 19) = 2.55$, $MSE = 25.17$, $p > .12$, replicating Kemp et al.'s (1990) finding, and indicating that contrast reversal and inversion have independent effects on identification of faces. *A priori* t-tests also showed that there was significant contrast reversal effect for upright faces considered alone, $t(19) = 10.47$, $p < .001$, and for inverted faces considered alone, $t(19) = 7.49$, $p < .001$.

For dogs, results did not show the same patterns as for faces with respect to either contrast reversal or inversion. The contrast reversal effect followed the same trend as that for faces, but the differences were much smaller (see Figure 4.5b). There was also no inversion effect. A 2-way ANOVA revealed no main effect of orientation, $F(1, 19) = 1.00$, $MSE = 47.32$, $p > .3$ and no interaction, $F < 1$, $MSE = 16.91$. There was a main effect of contrast, $F(1, 19) = 21.474$, $MSE = 17.83$, $p < .001$, and the difference between both original and both reversed was significant both for upright dogs, $t(19) = 4.07$, $p < .01$, and for inverted dogs, $t(19) = 2.91$, $p < .01$.

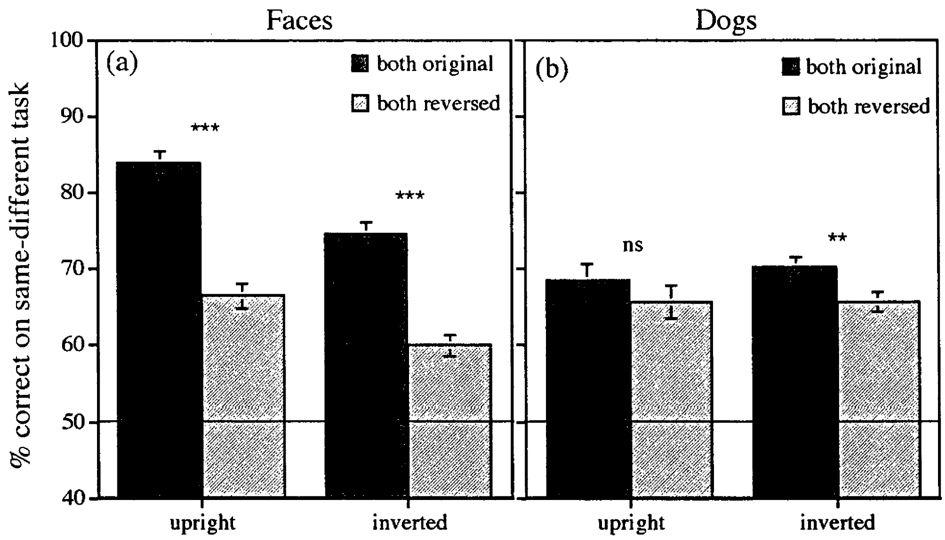


Figure 4.5. Experiment 5: Accuracy to say that both original contrast or both reversed contrast pairs are the same or different identity for (a) faces and (b) dogs. Error bars are equivalent to ± 1 SEM for the within subjects comparison of the two contrast conditions. ***, $p < .001$, ** $p < .01$, ns = $p > .05$.

Finally, results of the one-original-one-reversed condition are considered. As can be seen in Table 4.1, these pairs produced slightly less accurate identity judgments than the both reversed pairs. This effect occurred across the board. I attribute this to the additional disruption arising from the fact that pair members were never physically identical.

4.4.3 Experiment 5 - Discussion

For faces, the current results replicated several previous findings. Contrast reversal and orientation inversion both impaired identity-match judgements, and these effects were independent and additive (i.e, contrast reversal effects were as large for inverted faces as for upright faces; cf. Kemp et al., 1990). In agreement with previous authors, I suggest that the two effects arise from different stages of the visual processing stream. Presumably, the contrast reversal effects arise from difficulty in extracting shape-from-shading information in mid-level vision, while the inversion effects arise largely from disruption of holistic processing in high-level face processing (Hole et al., 1999; Kemp et al., 1990; Lewis & Johnston, 1997).

The new test was of contrast reversal effects for dogs. Despite the greater potential value of shape-from-shading information for labradors than for previously-tested object classes (particularly chairs and greebles), contrast reversal effects were still much weaker than for faces. Still, unlike in earlier results, a significant contrast reversal effect was found. The current results thus argue that, across a broader range of objects, contrast reversal effects can be merely disproportionately large in faces, rather than necessarily absent in objects.

A final observation of interest was the total lack of any inversion effect for dogs in Experiment 5. This result in the identity matching task contrasts with the small but nonzero inversion effect obtained on recognition memory (Experiment 4). A probable reason for this is that in the recognition memory task, each stimulus was seen only once whereas, in the present experiment, each dog stimulus was used many times. Inversion effects on object recognition are known to disappear rapidly with practice (e.g., Jolicoeur, 1985), consistent with rapid learning of part-based processing in noncanonical orientations. The stability of large inversion effects for whole faces in Experiment 5, however, is consistent with earlier findings that the inversion effect on holistic processing in faces is not removed even with substantial repetition (Chapter 3; McKone, 2004).

4.5 Experiment 6 – Composite Effect

Young et al.'s (1987) aligned vs. unaligned composite effect was used to provide a direct measure of holistic processing. The only object class previously tested in this task is greebles. Interestingly, greebles produced no composite effect at all in novices, rather than merely an effect that was smaller than that for faces (Gauthier & Tarr, 2002). In Experiment 6, I assessed whether this result would extend to the natural class of dogs (labradors).

The logic behind the aligned vs. unaligned version of the composite effect is that when two half faces of different individuals (e.g., the top half of George Bush with the bottom half of Tony Blair) are physically aligned, the two integrate to form a percept of a new individual, making it difficult to process the identity of a single half. When the same two halves are presented horizontally offset (unaligned), however, cues from early vision tell the face system that two people rather than one are being presented, thus allowing independent processing of each half to occur. Importantly, the aligned and unaligned conditions are otherwise comparable, in that simple response competition between the names ("Bush" arising from one half and "Blair" from the other) is the same in both cases. Thus, a finding that it is more difficult to name one half of the face in the aligned condition than in the unaligned condition provides clear evidence for holistic integration of the two halves in the former case. In Young et al.'s (1987) experiments, this composite effect was obtained for upright faces and not inverted faces (also see Carey & Diamond, 1994).

In Young et al. (1987), all stimuli were famous faces and so naming could be used as the response task. In the present study, however, the stimuli were novel. A technique suitable for novel faces was developed by Hole (1994; also see Le Grand et al., 2004). Hole simultaneously presented pairs of composite faces as shown in Figure 4.6, and subjects made a same-different identity response to one half (in his experiment, only the forehead). Each composite was formed from two different individuals. Across the pair of composites, the half-to-compare was either the same in identity or different in identity, and the half-to-be-ignored was always different in identity. Hole also jittered the position of each stimulus pair on each trial so that subjects could not pre-focus an attentional window on the expected location of the target half. Any such pre-focusing of an attentional "spotlight" would work against the emergence of a composite effect, as it could allow tuning out of the half-to-ignore at a very early stage of visual processing such that the face-recognition system never received information from that half.

Experiment 6 tested upright and inverted dogs as well as upright and inverted faces. Several features of Hole's (1994) procedure were incorporated. However, I used the full aligned versus unaligned comparison, rather than relying on upright vs. inverted differences in the aligned condition alone as did Hole. I also tested both foreheads and chins as target halves and, rather than measure reaction time as the dependent variable, used a limited presentation time and measured accuracy.

I considered only the results of same trials, and not different trials, to be of theoretical interest (see also, Hole et al., 1999; Le Grand et al., 2004). The predictions corresponding to the existence of holistic processing in each case are not the same. For same trials, the prediction is clear. The pattern of aligned versus unaligned difference should match that found in Young et al. (1987), namely the aligned condition should be less accurate than the unaligned condition. This is because when the top half of the same individual is paired with two different bottom halves, any holistic integration between aligned halves will make the tops appear less similar to each other than they really are, making it harder to say "same". Readers should be able to observe this perceptual phenomenon for faces in the upright orientation in Figure 4.6: the two top halves are the same identity (although differing in size and brightness), but this is more difficult to see in the aligned condition than in the unaligned condition.

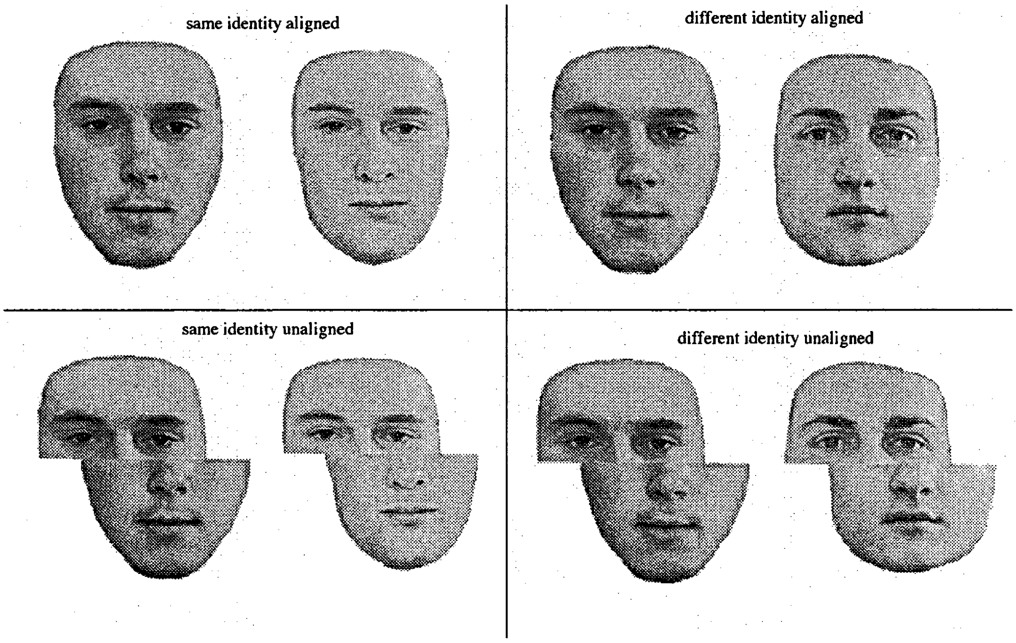


Figure 4.6. Experiment 6: Examples of same aligned, different aligned, same unaligned and different unaligned composite face pairs. In the examples here the half-to-compare is the forehead (so the chins are always different).

Now consider different trials. Here, the prediction corresponding to holistic processing is intrinsically unclear. When responding to the top half, if the two different bottom halves are quite different in appearance then perceptual integration should make the aligned condition more accurate than the unaligned condition (i.e. the reverse prediction to same trials). This is because the additional perceived dissimilarity should enhance performance in the aligned condition by making it easier to make the correct "different" judgement. However, it is also possible that two randomly-chosen bottom halves might happen to be fairly similar. In this case, perceptual integration with the top halves might make the two different top halves appear less different than they would by themselves. This would predict it should be harder to make the correct "different" judgement if holistic processing has occurred, rather than easier. The fact that holistic processing can thus predict either pattern of outcome for different trials (aligned > unaligned, or unaligned > aligned) means that analysing results of these trials is of no value in assessing holistic processing.

For <u>same</u> trials, I expected that the pattern corresponding to holistic processing would be revealed for upright faces, but not inverted faces. The question of interest was whether there would be any composite effect for dogs (see Figure 4.7 for example stimuli).

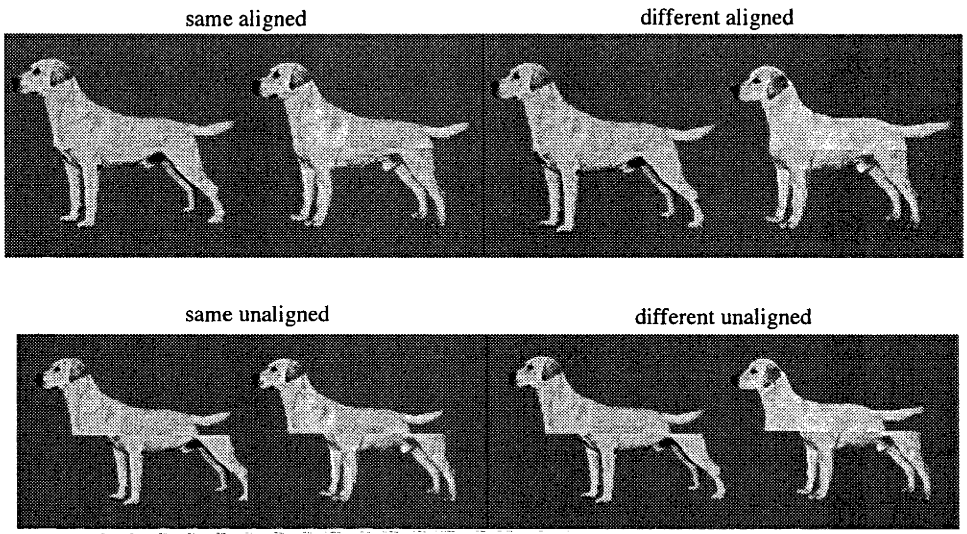


Figure 4.7. Examples of same aligned, different aligned, same unaligned, and different unaligned composite dog pairs. In each case the half-to-compare is the head/tail.

4.5.1 Experiment 6 - Method

4.5.1.1 Subjects for Experiment 6a (dogs only).

Twenty-four subjects completed Experiment 6a, half for course credit and half for $10 for the 1 hr session. None had participated in earlier experiments. Ages ranged from 18-43 (most 18-21); seven were male; all were Caucasian. All subjects reported normal or corrected-to-normal vision. These subjects completed the full experiment (faces and dogs), but a mistake in the creation of some of the face stimuli was discovered after the testing. Thus, only their data for dogs was used.

4.5.1.2 Subjects for Experiment 6b (faces and dogs).

Twenty-three additional subjects were tested on corrected stimuli for both faces and dogs, participating for course credit (N=8) or $10 (N=15). Ages ranged from 18-30; approximately half were male; all were Caucasian. All subjects reported normal or corrected-to-normal vision. Thirteen of these subjects had previously participated in Experiment 5.

4.5.1.3 Design.

A simultaneous presentation, same-different version of Young et al.'s (1987) composite paradigm was used. Each composite stimulus was created by taking the top half of one face (dog) and combining it with the bottom half of a different face (dog). For aligned composites, the two halves were aligned. For unaligned composites, the two halves were offset horizontally by approximately a quarter of the width of the face (dog). On each trial, two composites were presented simultaneously, either both aligned or both unaligned. The subject's task was to indicate whether a given half of the two stimuli (e.g., the top half) was the same or different in identity. The half-to-be-ignored (e.g., the bottom half) was always different for the two stimuli. Figure 4.6 shows examples of aligned same, unaligned same, aligned different, and unaligned different trials for faces. Figure 4.7 does the same for dogs. In instructions to subjects, halves were referred to as “forehead” versus “chin” for faces, and “head” versus “legs” for dogs, to avoid confusion with “top” and “bottom” when the stimuli were presented inverted.

Class (faces vs. dogs), orientation (upright vs. inverted), half-to-compare (forehead vs. chin for faces, head vs. legs for dogs) and alignment (aligned vs. unaligned) were all manipulated within subjects. Class, orientation and half-to-compare were blocked; for example, one block of trials presented dogs inverted for matching of the legs half. Order of testing for the blocked conditions was counterbalanced across subjects. Each block included an equal number of aligned and unaligned trials, presented in a different random order for each subject. On each trial, the composite pairs were presented for 600 ms, and accuracy of same-different responses was the dependent measure. Instructions were to judge identity with any changes in size or brightness/contrast to be ignored.

4.5.1.4 Stimuli.

Twelve dogs and twelve faces were chosen from the larger set of faces and dogs used for Experiment 4. Each class was divided into two sets of six exemplars. A given subject saw one set of exemplars for matching top halves and the other for matching bottom halves, with assignment of sets to half-to-compare condition counterbalanced across subjects. Stimuli within each set of six exemplars were chosen to satisfy several criteria: (1) they were all the same sex (all male); (2) they were chosen from only one book/database so that quality of photographs was similar across the set; and (3) when cut in half, any half-exemplar could be made to fit with any other half-exemplar reasonably well, allowing for changes in size and brightness/contrast only, with no distortion of the shape.

Forming composites. To make halves, each face was cut just below the eyes, and each dog was cut from mid-chest to just below the tail. Within a set of six exemplars, 30 composites were then formed by combining the top half of each exemplar with the bottom half of all other exemplars (6 top halves each combined with 5 possible bottom halves = 30 composites). Two versions of each of the 30 composites were created. In one version, the two halves were made to join up reasonably neatly by keeping the bottom halves unaltered and adjusting the size and brightness/contrast of the top half. These top-half-altered versions were used in blocks requiring matching of the top half, forcing subjects to base their decisions regarding "same" or "different" for a pair of composites on whether the top half had the same identity, rather than simply matching low level visual information such as size or brightness/contrast. In the other version of the 30 composites, the two halves were made to join up by keeping the top halves unaltered and adjusting the size and contrast of the bottom half. These bottom-half-altered versions were used in blocks requiring matching of the bottom half.

To manipulate alignment, the 30 composites of each version that had been created in the aligned format were then also created in the unaligned format. Each of the 30 composites was created in a left-offset format (i.e. the bottom half shifted to the left) and a right-offset format (i.e., the bottom half shifted to the right).

Forming pairs of composites. To form pairs of composites for simultaneous presentation, the following procedure was used. For each set of 30 composites (derived from 6 exemplars), all 60 possible same pairs were created (for a top-half-to-compare block these include: top1-bottom2 composite paired with top1-bottom3 composite; top1-bottom2 composite paired with top1-bottom4 composite; etc). All possible different pairs were also created. With the constraints that the top half must be different across the pair, the bottom half must be different across the pair, and that the two halves of a given original face should not appear on the screen simultaneously (i.e., as the top half of one composite and the bottom half of the other), there exist 180 such combinations for each set of 30 composites.

Structure within each block. Subjects performed one block of 240 trials for each class x orientation x half-to-compare condition (e.g. dogs inverted legs-to-compare). Within a block, each subject saw the 60 same trials for the assigned set of exemplars, plus 60 different trials selected randomly from the 180 possible such trials. These were all in the aligned format. A further 60 same unaligned trials and 60 different unaligned trials were presented, for the total of 240 trials per block. For every aligned composite trial, there was another trial that presented exactly the same composite items in the unaligned format. For unaligned trials, offset was to the left on half the trials and to the right on half the trials. The two composites shown always had offsets in the same direction, with the particular items assigned to each offset direction chosen randomly for each subject. Once a given subject's trial assignment had been determined for a given class and half-to-compare in the upright orientation (e.g., for dogs upright legs-to-compare), the trial structure of the inverted orientation was exactly matched to this. That is, all stimuli were exactly the same except rotated by 180° (and presented in a new random order).

Screen appearance. So that subjects did not forget which half to compare part-way through a block, all trials showed two short horizontal lines either above the stimuli (match the part on the upper half of the screen) or below the stimuli (match the part on the lower half of the screen). These half-to-compare indicator lines were set to the side so as to be visible in peripheral vision but not to interfere with processing of the composites.

The position of each composite on the screen was jittered from trial to trial. Because half-to-compare was blocked, jittering was particularly important to avoid pre-focusing of an attentional "spotlight" on the region where one half would appear. Each composite stimulus of a pair appeared in one of 5 positions. The horizontal displacement between the two composites (centre to centre) varied from 8 cm (10.2°) to 11 cm (13.9°) for faces and 7.5 cm (9.5°) to 8 cm (10.2°) for dogs. The vertical displacement varied from 2 mm (0.3°) to 1 cm (1.3°) for faces and 2 mm (0.3°) to 5 mm (0.6°) for dogs. Position selected was chosen randomly for each composite on each trial, with the constraint that the two were never presented at the same height.

4.5.1.5 Procedure.

Apparatus was as for Experiment 5. The task was explained to subjects using pairs of composite faces or dogs made from exemplars seen nowhere else in the experiment. Detailed instructions were given for the first class of stimuli (faces or dogs) to be seen by each subject. Instructions included a step by step talk through of a same trial and a different trial, to ensure the subject understood that the task was to match identity of the target half, not low-level appearance. Subjects were then warned that the stimuli on each trial would be available for only a brief period, and were then given 14 practice trials.

Subjects responded same via the green/right button and different via the red/left button on the button box. On each trial, the pair of composites was presented for a maximum of 600 ms followed by a blank screen until the subject responded. If the subject responded before 600 ms (this rarely happened), the stimulus was removed. The next trial began 200-400 ms after response. Subjects were informed that the correct answer would be same on 50% of trials and different on 50% of trials (pilot testing had revealed a bias to respond same as the default answer if no difference was found in the 600 ms). Prior to each experimental block, subjects were informed of the upcoming class, orientation and half-to-compare, and shown a schematic picture (e.g. an upside down dog with arrows pointing to the legs).

4.5.2 Experiment 6 - Results

Accuracy in matching halves (percentage of "same" responses for same trials, and percentage of "different" responses for different trials) was calculated for each subject for each condition[3]. This was done collapsing over top-half trials and bottom-half trials to give maximum power. (Results did not change when each half was analysed separately.) Mean accuracy for Experiment 6a (dogs only) and Experiment 6b (dogs and faces) is presented in Table 4.2.

---

[3] RTs are not reported because, where a brief presentation procedure is used, RTs are not meaningful. This is because, for example, if the subject is unsure of the correct response at the offset of the 600ms stimulus presentation, they could respond by guessing very quickly because no more stimulus information will be made available (i.e., very short RTs), or respond by considering their decision at length, perhaps attempting to remember the stimulus (i.e., very long RTs). In general, RT is only a meaningful measure when the stimulus stays on the screen until response.

To contrast dogs and faces within a single design, the results of Experiment 6b were first analysed. One observation from Table 4.2 was that there was slightly more bias to say that the dogs were the same, than to say faces were the same. This means that dogs had slightly (but significantly) higher accuracy than faces for same identity trials, $F(1, 22) = 4.89$, $MSE = 131.06$, $p < .05$, but that faces had a slightly, but not significantly, higher accuracy for different identity trials , $F(1, 22) = 3.72$, $MSE = 627.63$, $p > .06$.

Table 4.2. Mean accuracy (% correct) to match one half in the composite task of Experiments 6a and 6b. Part (a) shows same identity trials, part (b) shows different identity trials. SEM shown in brackets.

	N	upright aligned	upright unaligned	inverted aligned	inverted unaligned
a) same identity trials					
E6a - dogs	24	85.45 (1.91)	86.25 (1.85)	82.01 (2.36)	81.56 (2.55)
E6b – dogs	23	87.10 (2.71)	86.30 (3.33)	88.62 (2.07)	87.79 (2.18)
E6b - faces	23	81.67 (2.63)	85.83 (2.74)	83.51 (3.28)	83.88 (3.44)
b) different identity trials					
E6a - dogs	24	69.37 (3.44)	69.76 (3.33)	66.11 (2.95)	65.45 (3.20)
E6b – dogs	23	61.09 (4.04)	64.60 (4.10)	56.88 (4.83)	57.50 (4.71)
E6b - faces	23	74.85 (3.89)	69.75 (3.91)	62.97 (3.99)	61.01 (4.10)

As discussed earlier, the analysis of the composite effect considered same trials only. Results are plotted in Figure 4.8 (face at the top and dogs bottom left). A three-way ANOVA including class, orientation, and alignment condition revealed a significant 3-way interaction $F(1, 22) = 6.15$, $MSE = 6.64$, $p < .05$, indicating that alignment and orientation had different effects for faces and for dogs.

For faces, a two-way ANOVA then revealed exactly the expected pattern of results. There was a significant interaction between alignment and orientation, $F(1, 22) = 30.867$, $MSE = 2.70$, $p < .001$, and follow-up t-tests showed that this reflected a significant composite effect (i.e., aligned worse than unaligned) for upright faces, $t(22) = 4.85$, $p < .001$, but no composite effect (i.e., aligned equal to unaligned) for inverted faces, $t < 1$.

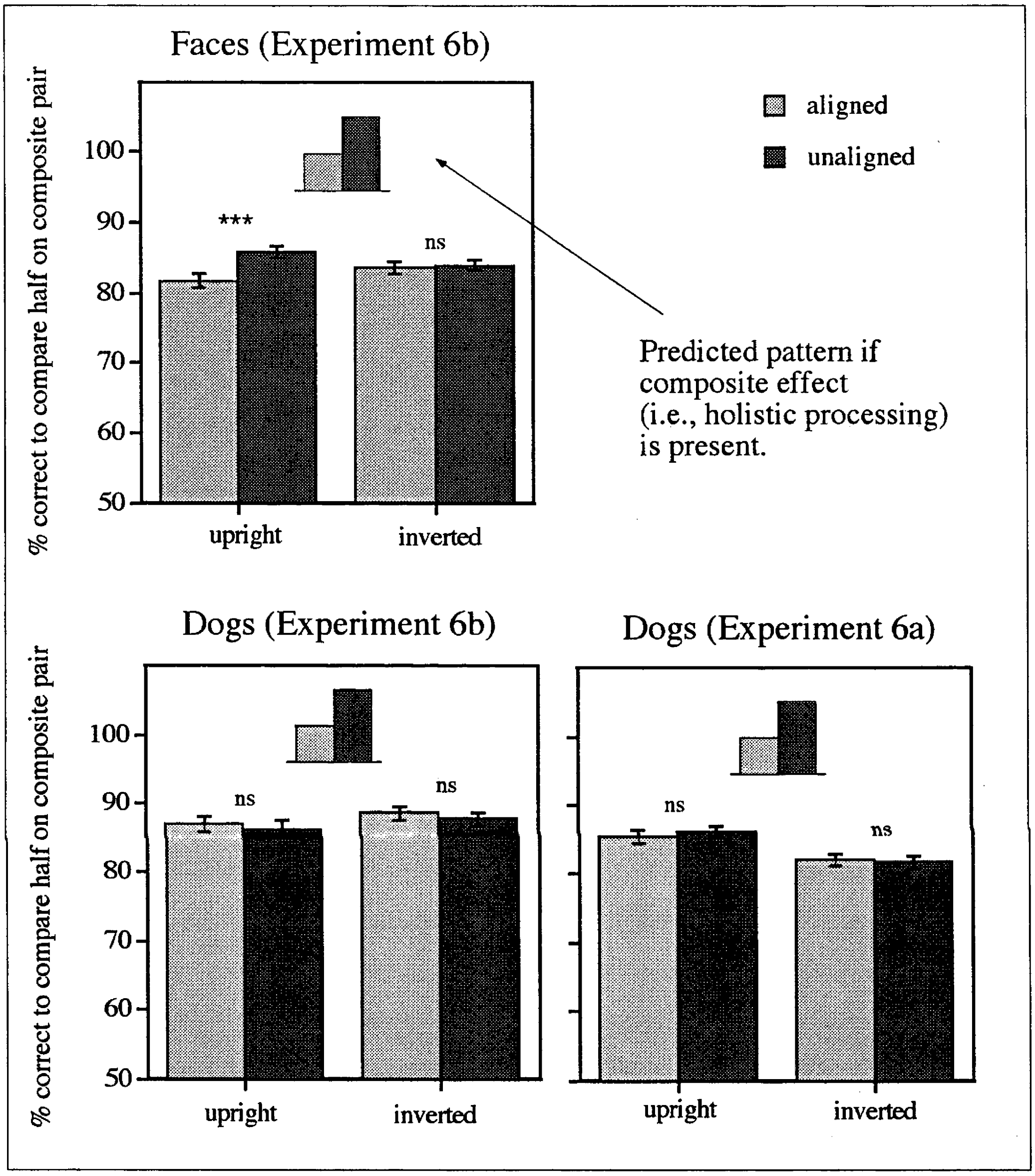


Figure 4.8. Experiment 6: Accuracy to compare one half of a pair of composites (simultaneous presentation same-different task). Results are for "same" trials only collapsed across top-half-to-compare and bottom-half-to-compare trials. Error bars are equivalent to ± 1 SEM for the within subjects comparison of aligned and unaligned conditions. *** $p < .001$, ** $p < .01$, ns = $p > .05$. The predicted pattern for a composite effect (holistic processing) is that accuracy should be lower for aligned than unaligned trials.

Thus, for upright faces, the aligned condition produced perceptual integration of halves into new wholes, making it hard to see that two top halves (for example) were the same when the respective bottom halves to which they were joined were different.

Turning to dogs in Experiment 6b, a different pattern emerged. There were no main effects of orientation, $F(1, 22) = 1.36$, $MSE = 38.11$, $p > .25$, or alignment, $F < 1$, $MSE = 15.28$, and no alignment x orientation interaction, $F < 1$, $MSE = 8.78$. *A priori* t-tests also revealed no differences between the aligned and unaligned conditions for either upright dogs, $t < 1$, or inverted dogs, $t < 1$, indicating no evidence of a composite effect in either case. The results of Experiment 6a (dogs only; Figure 4.8) agreed. There was no alignment x orientation interaction, $F <1$, $MSE=10.9$, and most importantly, there was no evidence of any composite effect for upright dogs $t < 1$. This independent replication of the null effect indicates there is no reason to doubt its reliability.

4.5.3 Experiment 6 - Discussion

The results showed a strong composite effect for upright faces, but no composite effect for inverted faces, replicating previous findings (Carey & Diamond, 1994; Hole, 1994; Hole et al., 1999; Young et al., 1987; Le Grand et al., 2004). For dogs, there was no composite effect in any condition. This provides the first generalisation of the previous result with greebles (Gauthier & Tarr, 2002) to a natural stimulus class.

Another important aspect of the results is that the composite effect for dogs was not merely smaller than that for faces but was in fact not there at all (aligned vs. unaligned difference = +0.8 Experiment 6a, –0.8 Experiment 6b, where a negative number indicates an effect in the reverse direction for a composite effect). Gauthier and Tarr obtained a similar result for greebles (aligned vs. unaligned difference = –2% change in reaction times). It thus seems that face-specific processing in the composite paradigm is associated with the presence versus absence of the signature phenomenon, rather than merely a disproportionately large effect for faces as in the inversion effect and the part-whole effect. This argues that the composite task provides a more powerful and direct way to assess holistic processing than other techniques.

4.6 General Discussion

The current experiments took three tasks known to produce strong signature phenomena for faces, and used these tasks with a natural stimulus class (dogs). My interest was in (a) whether the signature effects would be smaller for dogs than for faces, and (b) whether any effects might be completely absent for dogs.

Experiment 4 replicated many similar studies, in confirming that the inversion decrement on memory for labadors was much smaller than for faces, but was still significantly greater than zero. Thus, inversion effects on memory were merely <u>disproportionately large</u> for faces, rather than absent for dogs. In Experiment 5, I found that the contrast reversal decrement on identity-matching was much smaller for labradors than for faces, extending similar findings with manmade objects (chairs, blobs, greebles) to a natural object class. Unlike the earlier studies, however, dogs produced a small but significant contrast reversal decrement. Thus, the contrast reversal effect was also merely <u>disproportionately large</u> for faces. In Experiment 6, I tested Young et al.'s (1987) aligned vs.unaligned composite effect. Here, I obtained a result where the signature phenomenon was <u>absent for objects</u>. Despite establishing a clear composite effect for faces, there was no effect for labradors, extending a similar previous finding with greebles to a natural object class.

4.6.1 <u>How reliable is the "zero" effect in the composite task?</u>

In the final experiment, an important issue is the reliability of the conclusion that the effect was completely absent for dogs. I address this issue, first, by considering the statistical reliability of the data. In the composite task of Experiment 6, each subject completed a relatively large number of trials (60 for every condition), and I ran the dog section of the experiment on two independent sets of subjects (N=24 and N=23). Each set of subjects produced only a tiny accuracy difference between the aligned and unaligned conditions for dogs. This difference went in opposite directions across tests (+0.8 % correct in one case, –0.8 % correct in the other) and, collapsed across the total of 47 subjects, the 95% confidence interval on the size of the composite effect was tightly centred around zero (-1.44 to +1.47 % correct) and allowed only a trivially small positive effect. These results argue that there was genuinely no composite effect.

An equally important factor in evaluating a "zero effect" conclusion is the nature of the object class chosen. I consider the findings to be particularly strong because of my careful matching of face and non-face stimuli. While there is no one object class that perfectly matches faces on all possible variables, dogs provide a good choice. Ways in which dogs are matched to faces include having a canonical upright, a shared first-order configuration, some part boundaries that are fuzzy rather than sharp, and a moderate degree of surface texture. These properties are not true of all other objects; for example, houses and greebles include clear part boundaries, and thus potentially could be more separable into parts than dogs. Moreover, dogs are natural stimuli produced by genetic mechanisms, meaning that individuals differ from each other on a great many shape dimensions at once, and that, as for faces, these shape differences are distributed all over the spatial extent of the dog. These properties potentially require a greater emphasis on global processing than for object classes where exemplars differ on only a few dimensions, as is common in manmade and artificial objects.

4.6.2 Which tests are best?

An active debate in the literature is whether apparently face-specific cognitive phenomena might emerge for objects under certain conditions. The most common theory is the "expertise hypothesis", which predicts that objects should show face-like processing where the subject has extensive expertise in making individual-level discriminations (e.g., Diamond & Carey, 1986; Gauthier & Tarr, 2002; but see McKone & Kanwisher, in press); other ideas are that within-class discrimination alone might be sufficient to produce face-like processing (e.g. Damasio, Damasio, & Van Hoesen, 1982), or that some objects are processed like faces when right-hemisphere coordinate relations rather than left-hemisphere categorical relations are required (Cooper & Brooks, 2004). To test any of these various theories, it is important to choose tasks that are known to reliably dissociate faces from objects under ordinary circumstances. I now summarise my views on the status of various tasks that could be used as measures of face-specificity.

A first point is that, if the intention is to assess holistic face-like processing, then it is essential to use a task that has a demonstrated association to holistic processing in faces, such as producing its signature phenomenon for upright faces, but not for inverted faces, scrambled faces and/or isolated face features. The inversion effect on memory, the part-whole task, the composite effect and the peripheral inversion paradigm (McKone, 2004) all meet this criterion. The contrast reversal effect does not. (This does not mean that it is of no value to assess contrast reversal effects in experts, merely that any emergence of face-like effects for objects would not be indicative of holistic processing.)

Second, there is the question of the degree of dissociation of processing between faces and objects. It is clear that inversion effects on memory and the part-whole effect do not totally dissociate face from object processing. Both tasks produce signature effects that are much larger for faces than for objects, but which are still present to some extent for objects. These results argue that neither of these tasks provides a pure measure of face-specific processing. Instead, in both cases, it seems likely that the task partly taps face-specific holistic processes, but also includes some general component. For the inversion effect, this general component presumably is the advantage to memory provided by the familiar structure available in the canonical upright (Maurer et al., 2002). In the case of the part-whole effect, the general component is presumably some kind of context, gestalt or global processing effect analogous to the word superiority effect (Davidoff & Donnelly, 1990).

Turning to contrast reversal, the results of earlier studies had suggested that contrast reversal effects were absent for objects. No significant decrement had been reported for three classes of manmade objects, namely greebles, chairs, and "blobs" (Gauthier et al., 1998; Nederhouse et al., 2002; Subramaniam & Biederman, 1997). My results, however, demonstrated a small but nonzero contrast reversal effect for dogs. I therefore conclude that, looking across a broader range of object types than previously tested, contrast reversal effects also do not purely dissociate face and object processing.

Conversely, the composite paradigm appears to provide a complete dissociation, and also provides a theoretically strong measure of holistic processing. The composite effect has been found to be absent for both greebles and dogs, thus spanning artificial and natural object classes.

I conclude that the composite effect is the best of the standard tasks to use when assessing whether "face-like" holistic processing emerges for objects under special circumstances, such as when the subject is an expert in the relevant domain. This conclusion is important because this is exactly the task that has <u>not</u> been substantially investigated in experts. Instead inversion effects (e.g, Diamond & Carey, 1986), and part-whole effects have been tested (Gauthier & Tarr, 1997; Tanaka et al. 1996, cited in Tanaka & Gauthier, 1997).

4.6.3 Implications of the results for the "within-class discrimination" and "coordinate representations" hypotheses

I finish with a brief comment on the implications of the results for two theories (other than expertise) of why faces are special. One older view is the within-class discrimination hypothesis (e.g., see Faust, 1955, cited in Benton, 1990). This view suggested that faces might merely appear to be special because, in most everyday circumstances and some experimental tests (e.g., early studies of prosopagnosia), individual level discrimination of faces (Bill vs. Sam) had been contrasted with basic level recognition of objects (chair vs. table) rather than individual level discrimination of objects (labrador 1 vs. labrador 2). This hypothesis is generally held to have been disproved and, indeed, all the studies reviewed in this chapter, and more fully in Chapter 1 (inversion effects, part-whole effect, and so on) have contrasted individual processing of faces with individual processing of objects, and confirmed different effects for faces and objects in every case. Despite this evidence, however, the within-class discrimination hypothesis has been considered in a number of surprisingly recent papers (de Gelder & Rouw, 2000; Gauthier, Tarr et al., 2000; Tarr, 2003; Tarr & Cheng, 2003). Note that, in choosing dogs as a well matched stimulus to faces, and requiring individual-level identification in each case, the current results confirm on three separate tasks that the within-class discrimination hypothesis cannot explain "special" processing for faces.

Another theory comes from the literature on object recognition. Cooper and Brooks (2004) have suggested that, while normal face processing requires "coordinate relations" (detailed metric information), much object recognition requires only "categorical relations" between parts (above, below, side-of, etc). These authors then argue that processing of objects relies on coordinate relations – that is, the same type as those supporting face recognition – when the objects to be distinguished are all reasonably similar in form and detailed metric information about shape becomes necessary to distinguish between them (see also White, 2002). However, the objects required to be discriminated in the present experiments (individual labradors) are all extremely similar in basic shape, and yet showed patterns of data that were very different from those for faces. Thus, while it might well be that the categorical versus coordinate relations distinction is useful for understanding object processing (e.g., Hellige & Michimata, 1989; Keane, Hayward, & Burke, 2003; Kosslyn et al., 1989), Cooper and Brooks' assumed equivalence between coordinate relations and the style of processing used for faces cannot be correct.

CHAPTER 5: EXPERTISE AND FACE-LIKE PROCESSING: THE CASE OF LABRADOR EXPERTS.

5.1 Overview

The idea that apparently "special" face processing might be due to greater expertise with faces compared to other objects was first tested by Diamond and Carey (1986). They showed an inversion effect for dog experts looking at dogs which was as large as that for the same experts looking at faces. Here I attempted to replicate this result for labrador retriever experts looking at labradors. Further, I tested for face-like effects on a test of contrast reversal, and most importantly tested for holistic/configural processing using Young et al.'s (1987) composite task. In no case did labrador experts show face-like effects for labradors. Experts showed a smaller inversion effect for dogs than for faces, no effect of contrast reversal, and no holistic processing. I conclude that face processing does not result from a generic expertise system, in which holistic processing can be learned with practice at any point throughout life. Instead interpretations in terms of either a critical period for developing face processing (Le Grand et al., 2001) and/or an innate representation of face structure are discussed.

5.2 Introduction

The previous chapter compared results for faces and labrador dogs in dog novices on three tests that had been suggested to produce face-specific effects. Results confirmed that within-class discrimination alone was not sufficient to induce face-like processing of objects on these tests. The current chapter investigates the more interesting possibility that within-class processing of objects-of-expertise would induce face-like processing in dog experts, using the same tests as in the previous chapter.

Given that people are generally much more expert at distinguishing individual faces than at distinguishing individuals from other classes of stimuli, the idea that differences between faces and objects might be due to expertise (Meadows, 1974) is the major competing theory to the claims of domain-specificity (i.e., that faces *per se* are special). Diamond and Carey (1986) suggested that any object might be holistically processed if three criteria are met. First, the objects of a class should share a first-order configuration; this is satisfied by dogs, for example, in that (nearly) all have a body, with four legs beneath, a head at one end and a tail at the other. Second, individual level identification is required; this is satisfied if the task requires discriminating, for example, labrador 1 from labrador 2. Third, sufficient experience at discriminating members of the class is required; this is satisfied by using subjects with many years' experience in the relevant domain (e.g., dog show judges if using dogs as the stimuli).

Currently, the question of whether expertise can change the style of processing used for objects is a matter of active debate. Some authors argue that "certain classes of objects might be 'special' not because of their intrinsic status, but because we have expertise with them" (Palmeri, Wong, & Gauthier, 2004, p. 378); or that "while face recognition is certainly the most complicated discrimination task most of us ever learn to perform, it is still part and parcel of general recognition mechanisms, albeit mechanisms that have been tuned to recognise specific faces through many years of experience with objects" (Tarr, 2003, p. 192). Conversely, other authors argue that "Substantial evidence supports the domain-specificity of face processing in humans" (McKone & Kanwisher, in press, p.339) or that "several lines of evidence suggest that [the] ability to individuate faces may result from the operations of a specialised neural module that encodes faces, and not other objects" (Rhodes, Byatt, Michie, & Puce, 2004, p189). My own opinion is that most of the evidence argues against the expertise hypothesis and in favour of domain-specificity.

Expertise has generally been tested at two levels of experience - laboratory trained (e.g., greeble experts) and real-world experts (e.g, car experts) – and with both manmade objects (e.g., greebles, cars) and natural objects (e.g, birds). It is important to note that to give the strongest argument either for or against the expertise hypothesis a wide range of objects needs to be considered, preferably combined with experience for that object class. I will begin with a review of the previous studies on expertise for objects. These include neuroimaging and neuropsychological studies assessing whether objects-of-expertise are processed in the same brain areas as faces, and behavioural evidence regarding whether the same style of cognitive processing is used for both objects-of-expertise and faces.

5.2.1 Neural processing of faces versus objects-of-expertise

5.2.1.1 Expertise and the FFA.

In novices, the Fusiform Face Area (FFA) is more activated for faces than a wide range of other objects (Grill-Spector, Knouf, & Kanwisher, 2004; Kanwisher, McDermott, & Chun, 1997; Rhodes, Byatt et al., 2004; Tong, Nakayama, Moscovitch, Weinrib, & Kanwisher, 2000). The expertise hypothesis, however, suggests that the FFA might be a general expertise area rather than a face specific area. If this were the case, it would predict that experts with a given object class (e.g., car experts) should show greater activation (percent signal change from baseline in the BOLD response) within the FFA for that class than shown by subjects with less expertise for that class (i.e., car novices). Further, if the level of expertise is high enough, activation for objects-of-expertise should approach that for faces.

In the first study to investigate this issue, Gauthier, Tarr, Anderson, Skudlarski, and Gore (1999) tested BOLD response in the FFA before and after seven hours training with greebles, using both a passive viewing and one-back matching task (the latter ensures attention to individual identity). They compared the activation difference between upright and inverted greebles, as well as the difference between upright and inverted faces. Gauthier et al. found that the difference between upright and inverted greebles was more similar to the difference between upright and inverted faces after greeble training than before, and interpret this as increased activation in the FFA with expertise. However, as McKone and Kanwisher (in press) point out, given that the FFA shows only a small inversion effect (Kanwisher, Tong, & Nakayama, 1998) this is a rather odd choice of measure. Also, unfortunately, Gauthier et al. did not report the direct percent signal change from baseline for faces and greebles, so it is not clear whether this study found an increase in activation to upright greebles with training (as opposed to a decrease to inverted greebles which would also increase the difference between upright and inverted), or whether activation for upright greebles ever approached that for upright faces.

Several other studies have reported more direct comparison to faces, testing car, bird, and Lepidoptera (butterfly and moth) experts. In these studies, real-world experts of three to thirty years' experience were tested. Gauthier, Skudlarski, Gore, and Anderson (2000) tested bird and car experts on a one-back matching task comparing activation for birds, cars and faces. Bird experts, who were also novices with cars, showed higher activation for birds than for cars in the FFA. Car experts, who were novices with birds, showed approximately equal activation for birds and cars. In both cases the activation for objects was less than the activation for faces. These results could be taken as indicating that bird experts, at least, use the FFA to process birds, with the weaker activation to birds than to faces reflecting less experience with birds than faces. However, Grill-Spector et al. (2004) showed that event-related activation[1] to birds was higher than other object classes even in non-experts (possibly because birds have heads), thus questioning whether the result found in Gauthier, Skudlarski et al. was due to expertise rather than reflecting an already stronger than usual response further enhanced by attention (Wojciulik, Kanwisher, & Driver, 1998).

Grill-Spector et al. (2004) also tested car experts. Experts showed no greater activation for cars than did novices. There was no correlation between success at identifying or detecting cars and FFA activation in car experts, thus failing to replicate Gauthier, Skudlarski et al.'s (2000) bird expert result in a more sensitive design. Rhodes, Byatt et al. (2004) found slightly stronger FFA activation for Lepidoptera in experts than novices, but activation for faces was also slightly stronger in the Lepidoptera experts than in novices (and in both groups activation for faces was stronger than activation for Lepidoptera). Further, when voxels active for faces were compared to those active for Lepidoptera the overlap was less than 15% even in experts. Thus, while there is still active controversy, the results from neuroimaging studies seem more consistent with the idea that faces *per se*, not expertise, leads to strong activation of the FFA.

5.2.1.2 Expertise, the N170 and the M170

The temporal properties of activation to faces versus objects-of-expertise have also been studied. Based on the face-specific N170 response in ERPs, Rossion, Gauthier, Goffaux, Tarr, & Crommelinck (2002) studied an N170 response in greeble experts. They argue that the delayed and enhanced response in inverted faces compared to that for upright faces is a more reliable measure of face-specificity than comparing the response to faces (or objects-of-expertise) to that for other objects. They therefore compared the difference in response to upright and inverted greebles after training to that for faces. They found that results for greebles after training were more similar to

[1] Correlation on a trial-by-trial basis between the stimulus and the BOLD response.

those of faces than results before training, being both increased in amplitude and delayed in latency. However, this effect was only found in the left hemisphere, whereas faces usually show a bias to right hemisphere processing (e.g., Gauthier, Skudlarski et al., 2000; Rhodes, 1993).

Using real-world experts, Gauthier, Curran, Curby and Collins (2003) found a difference in N170 between car experts looking at cars and novices looking at cars. Car experts showed larger amplitude in the N170 to cars than did novices but, as in Rossion et al. (2002), this was only in the left hemisphere. Comparing bird and dog experts, Tanaka and Curran (2001) reported a larger N170 response to objects-of-expertise than to the other class. However, there was no comparison made to faces, meaning that the size of effect relative to faces cannot be assessed. Given that there is some N170 effect for many object classes (Rossion, Gauthier et al, 2000) this makes it difficult to say whether the N170 is related to expertise rather than faceness. Further, as Tanaka and Curran themselves note, the site at which this N170 was found was again different to the site at which it is usually found for faces (although in this case it was in the right hemisphere). Thus, these three studies suggest that an N170 can be identified which is larger (and delayed) for experts looking at objects-of-expertise than for novices on the same objects. However, this appears to be coming from different brain areas than the N170 for faces (e.g., left versus right hemisphere).

Further, using MEG, Xu, Liu, and Kanwisher (2004) showed that the M170 did not correlate on a trial-by-trial basis with successful car identification in car experts, although it did correlate with successful face identification. Overall, the evidence from the N170 and M170 seems to support the domain specificity of face processing rather than the expertise hypothesis.

5.2.1.3 Expertise and neuropsychology.

Results from neuropsychology also indicate face specificity rather than expertise. If faces are special due to expertise, then damaging face-processing areas of the brain should also affect any expertise with objects. Similarly, intact face processing should mean that processing for objects-of-expertise also remains intact. That is, it should not be possible to dissociate expertise with objects and ability to process faces.

Patient CK is object agnosic, but not prosopagnosic, from a brain injury. Despite his extremely good face recognition skills, CK cannot recognise the individual toy soldiers or aeroplanes with which he was an expert before his injury (Moscovitch, Winocur, & Behrmann, 1997). There are also three cases of prosopagnosic subjects who retained, or gained, expertise with non-face objects after the onset of prosopagnosia. ELM was an expert with brass instruments before his stroke. Although after brain injury he was worse than controls at recognising faces (and individual stringed instruments), he was better than controls at recognising brass instruments (Dixon, Desmarais, Gojmerac, Schweizer, & Bub, 2002). Similarly, WJ learned to individually recognise his sheep after acquiring prosopagnosia (McNeil & Warrington, 1993), and Edward (a developmental prosopagnosic) did not differ from normal controls at learning to recognise greebles (Duchaine, Dingle, Butterworth, & Nakayama, 2004). These studies thus suggest a double-dissociation between the area of the brain used for face-recognition and that used for recognition of objects-of-expertise.

5.2.2 Behavioural studies of objects-of-expertise

Behavioural studies have tested experts in several paradigms which produce effects "special' to faces in novices. The next sections will review studies on effects associated with holistic processing (inversion, part-whole, composite) as well as contrast reversal. Results of these tasks were reviewed for novices in Chapter 4. The question of interest here is whether the results are different for experts; in particular, whether effects are bigger for objects-of-expertise and whether these ever reach the same size as effects for faces.

5.2.2.1 Expertise and inversion effects on recognition memory.

Inversion effects do not directly measure holistic processing (as previously noted). Indeed, given that most objects show some inversion effect, they are likely to be partly due to general familiarity with objects in one orientation. The expertise hypothesis explains disproportionate inversion effects for faces by noting that our level of experience with upright faces is much greater than our experience with inverted faces, whereas for other objects the level of experience at different orientations is more similar. Thus, becoming an expert with a particular object class in an upright orientation should increase the size of inversion effects. In terms of empirical tests, the effect of inversion for objects-of-expertise was first studied by Diamond and Carey (1986). They conducted two experiments with dog experts. In the first (Experiment 2 in their paper), the inversion effect for dog experts was not significantly larger than that for novices, and was still smaller than for faces (a 19% decrement for faces, 12% for dogs in dog experts; a 17% decrement for faces, 8% for dogs in dog novices). In this experiment the breed of dogs used as stimuli was not specifically matched to the breeds with which the experts were expert. When breed was more carefully matched to expertise (in Diamond & Carey's Experiment 3), dog experts showed the same sized inversion effect for faces and dogs (20% for faces, 22% for dogs). This effect for dogs was significantly larger than that for novices (less than 5%). This result was particularly convincing because upright performance for faces and dogs was the same for dog experts.

This result has been taken as strong evidence for the expertise hypothesis, and is regularly cited as such (e.g., Collishaw & Hole, 2002; Fallshore & Schooler, 1995; Morton & Johnson, 1991; Rhodes, 1993). However, the finding of a face-sized inversion effect has never been replicated. Bruyer and Crispeels (1992) claim to have replicated Diamond and Carey's result in handwriting experts. The abstract implies that they have replicated the face-sized inversion effect, but the actual paper reports replicating Diamond and Carey's Experiment 2 results. Specifically, experts showed a larger inversion effect for handwriting than novices (approximately 10% for experts vs. approximately 5% for novices), but the effect was still only half the size of that for faces (approximately 20% for both groups). It could perhaps be argued that this is not a fair comparison, as the upright performance was not matched for faces and handwriting. However, performance for inverted stimuli (for which experts and novices should be equally inexpert) was matched. In a similar example, Rossion et al. (2002) found a significant interaction between orientation and level of expertise for greebles. In a sequential same-different task the inversion effect on reaction time was 25 ms for greebles in greeble novices and 46 ms in greeble experts (accuracy was high and similar in all conditions), but this effect was still smaller than the 75 ms inversion effect for faces. (In this case performance for upright faces and greebles was well matched.)

Several other studies include information on inversion effects in objects-of-expertise for experts and novices, but with no comparison to faces. Gauthier, Williams, Tarr, & Tanaka (1998) compared upright greebles to "mis-oriented" greebles (data collapsed over greebles presented at 60°, 120° and 180° from upright) and, contrary to the expertise hypothesis, found a larger mis-orientation effect for novices than experts (231 ms vs. 180 ms, respectively). Gauthier, Skudlarski et al. (2000) compared the performance for bird and car experts looking at both birds and cars. While the car experts showed a larger inversion effect for cars than for birds (differences in the sensitivity measure d' was cars = 0.84, and birds = 0.05), the bird experts showed the same pattern (cars = 0.57, birds = 0.30). In this study performance was not matched across the two stimulus types at either orientation. Xu et al. (2004) also reported d' prime for novices and car experts looking at cars. Their article reports a larger inversion effect for experts than for novices on this measure. However, performance was again not matched for either orientation, and experts were in fact much better in both orientations than novices. Moreover, when results are reported as percentage correct (Diamond & Carey's, 1986, measure), Xu et al.'s data show no difference in the size of inversion effect between the two groups (upright: car experts = 90%, car novices = 71%; inverted: experts = 82%, novices = 63%; inversion effect = 8% for both groups; personal communication, Xu & Kanwisher, September, 2004).

Overall, results suggest that expertise increases inversion effects in some cases (e.g., handwriting experts in Bruyer & Crispeels, 1992; greeble experts in Rossion et al., 2002), but has no effect in others (e.g., greeble experts in Gauthier et al, 1998; bird experts in Gauthier, Skudlarski et al., 2000; car experts using the % measure in Xu et al., 2004). Moreover, except for the original study by Diamond and Carey (1986) there is no evidence that inversion effects are the same size as for faces. Given that face-specificity manifests as a disproportionate inversion effect, this is an important point.

Another problem for the expertise hypothesis is that in some studies expertise obtained for objects in the upright orientation improved experts' recognition of inverted objects with respect to that of novices. This is what occurs for both Xu et al.'s (2004), and Rossion et al's data (2002; a 7% and 81 ms improvement in experts for inverted greebles compared to a 2% and 102 ms improvement for upright greebles). It is not clear how the expertise hypothesis can explain this, while at the same time explaining the disproportionate inversion effects for faces.

5.2.2.2 Parts versus wholes paradigm.

As discussed in Chapter 4, in novices, the part-whole effect (Davidoff & Donnelly, 1990; Tanaka & Farah, 1993) is similar to the inversion effect in being disproportionately large for faces compared to objects, rather than being present for faces and absent for objects. Several studies have compared the part-whole effect for novices to that for experts with non-face objects. Most of these have been conducted using greebles and greeble experts and did not test faces as a comparison. For upright greebles, Gauthier and Tarr (1997) found that greeble experts were significantly more accurate at recognising a part in the learned whole greeble than in isolation (11% difference), there was a small non-significant effect for inverted greebles in the same direction (5% difference), and the overall pattern of results was not reliably different from that of greeble novices (part-whole difference for upright = 5%, for inverted = 2%). Gauthier and Tarr also tested a transformed version of the part-whole effect, in which a part is tested either in isolation, in the original whole or in the context of a whole with a spacing change between features (the "transformed" condition). The usual finding for faces (Tanaka & Sengco, 1997) is that performance is more accurate and/or faster in the original condition than in the transformed condition. Gauthier and Tarr (1997) found that for greeble experts there were no significant effects on accuracy, but parts were recognised significantly faster in studied whole greebles than in transformed greebles. For greeble novices, there were no differences between original and transformed greebles. The interaction between expertise and condition (transformed vs. original) for reaction time (RT) was not quite significant (the interaction for accuracy is not reported).

Gauthier et al. (1998) also tested both the original and transformed versions of the part-whole effect. For experts, one of the parts tested (the "quiff") was recognised significantly better in the original whole than the isolated part configuration, and nearly significantly better than the transformed configuration ($p=.059$). However, results for the other two parts trended in the opposite direction (worse performance for parts in the original whole configuration than alone). Moreover, for novices, recognition was also better for original whole greebles than for isolated parts for one of the three tested parts (the "dunth"). Thus, this study did not show any consistent effect of expertise.

Gauthier and Tarr (2002) found a slight advantage for parts in the original configuration over parts presented in isolation for both novices and experts (a difference of about 1.4 on a d' measure for both groups). There was also an advantage for recognising parts in original, as compared to transformed greebles, for both groups but this was primarily for recognising the particular part which had been moved to make the transformed stimuli (a difference in d' of about 1, compared to a difference in the other parts of 0.3). Overall, the greebles studies have found small part-whole effects that are no larger in experts than in novices. None of these studies tested faces as a comparison, so it is difficult to directly assess whether the effects for greebles would be of a similar size to those found for faces.

The only study to test the part-whole effect in real world experts is by Tanaka et al. (1996, cited in Tanaka & Gauthier, 1997). They tested car experts, dog experts and biological cell experts and compared the results for objects-of-expertise to those for faces. For experts, there were small part-whole effects for each group for objects-of-expertise (approximately 10% for cells, 6% for cars and 8% for dog faces). These effects were not larger than the effects for novices except perhaps for dog-faces (approximately 16% for cells, 8% for cars and 2% for dog faces). In all cases these were also smaller than the effects for faces (approximately 26% for cell experts, 18% for car experts and 20% for dog experts, and approximately 25% for object novices).

In summary, there is no evidence that expertise reliably increases the size of the part-whole effect for either greeble experts or real-world experts compared to the small effects found in novices. Gauthier and Tarr (2002) argue that the part-whole effect provides no evidence for face-specific processing because the increased effect for faces reflects only increased expertise. This is inconsistent with the findings above, namely that expertise with non-face objects does not increase the size of the part-whole effect compared to that for novices.

5.2.2.3 Composite effect.

As shown in Chapter 4, the composite effect (Young, Hellawell, & Hay, 1987), unlike the part-whole effect, is absent for novices looking at non-face objects. The composite effect has only been tested in one group of experts, that is, greeble experts. Using a direct implementation of Young et al's (1987) original procedure with famous faces, Gauthier et al. (1998) tested experts' (but not novices') ability to name one half of a familiar greeble, either aligned or unaligned with another half greeble. For accuracy there was no difference between aligned and unaligned conditions where the halves were from the same "family", and there was a trend for aligned composite greebles with halves from different "families" to be named more accurately than unaligned composites. This is the opposite pattern to that predicted for a composite effect. For RT, when the halves came from different families there was a trend for aligned composites to be named faster than unaligned (again, the opposite pattern to the standard composite effect), but for composites made from halves from the same family there was a trend for aligned composites to be named slower than unaligned composites. There are no statistics reported for these comparisons, but from an examination of the error bars it seems likely that the aligned and unaligned conditions do not differ significantly in any of these comparisons.

Gauthier and Tarr (2002) conducted a sequential same-different version of the composite effect with both greeble experts and novices. Results for aligned and unaligned same trials (the appropriate comparison for looking for a composite effect in this version of the composite task; see Chapter 4) are only reported for RT. Gauthier and Tarr claim that the composite effect increased with expertise. However, the close-to-significant interaction with expertise ($p = .07$) reflected a pattern in which there was initially no difference between processing speed of aligned and unaligned composites, a difference in the reverse-to-predicted direction (i.e., aligned faster than unaligned) in the second and third sessions of training, and then no difference again in the final two sessions. From their data (Gauthier & Tarr, 2002, Figure 4c, p. 439, Table 2, p. 441) it is clear that there is no difference between aligned and unaligned trials at the end of training (i.e., in experts).

In summary, in the one object class tested, there is no evidence of a composite effect for experts viewing objects-of-expertise. Note, however, that no natural object classes or real-world experts have been tested. It is thus still possible that extensive real-world experience with a natural object class might produce a composite effect with objects-of-expertise.

5.2.2.4 Contrast reversal.

Results from Chapter 4 showed that, in novices, there was a disproportionate contrast reversal effect for faces compared to labrador dogs. Previous research had also shown essentially no effect of contrast reversal for greebles, blobs or chairs in novices (Gauthier et al., 1998; Nederhouser, Mangini, Biederman, & Kazunori, 2002; Subramaniam & Biederman, 1997). Contrast reversal does not assess holistic processing; however, as noted in Chapter 4, it does seem to be a task on which faces and objects differ (at least in the magnitude of the effect) and it thus worth investigating in experts.

There have been two studies of the effect of contrast reversal on objects-of-expertise. In the only published study, Gauthier et al. (1998) found that greeble experts were worse at recognising contrast reversed greebles compared to original contrast greebles. They were also worse at recognising contrast reversed greebles than were greeble "novices". (The amount of training given to "novices" is not reported, but it was enough that they could verify the names of greebles at an accuracy not significantly different from experts.) Both groups were the same at recognising original contrast greebles. Reporting in a conference presentation, Nederhouser et al. (2002), found no effect of contrast reversal on matching blobs for either novices or experts (the only effect of expertise was to make experts faster at the task overall). Thus, there are contradictory results in the two studies of the effect of contrast reversal on objects-of-expertise. Further, natural objects and real-world experts have not been tested.

5.2.2.5 Summary of behavioural studies.

Inversion effects have been taken as strong evidence for the expertise hypothesis. However a review of the literature has shown face-size inversion effects for object-of-expertise only in Diamond and Carey's (1986) study with dog experts (and only in Experiment 3). Other studies reported smaller or no effects of expertise on inversion. Also, inversion effects do not directly demonstrate holistic processing (i.e., the aspect suggested to be special to faces). Two tests which do directly test holistic processing have been performed with experts. The part-whole effect, like the inversion effect, is smaller for objects than for faces in novices. For experts, there is no increase in the size of the effect for a range of natural and manmade objects. There is no composite effect for greeble experts looking at greebles (i.e., no holistic processing). However, no natural object classes or real-world experts have been tested on this paradigm. This is important as training in an experimental setting may be insufficient to develop this face-specific processing, and thus a finding of no effect in experiment trained experts cannot refute the expertise hypothesis.

5.2.3 Present study

The aim of the present study was to test real-world experts on the three tasks used in Chapter 4. These were a test of the effect of inversion on recognition memory, a test of the effects of contrast reversal (and inversion) on a simultaneous matching task, and a test of holistic processing using Young et al's (1987) composite effect. Although it would be very interesting if face-like effects were found after only ten hours of training, the fact that they are not does not necessarily mean that the expertise hypothesis is wrong. Some of the potentially most interesting tests of face-like processing (e.g., the composite effect) have not previously been tested with real-world experts.

The particular experts tested here were experienced with labrador retrievers. Labradors are considered to be particularly suitable stimuli to compare to faces for the reasons listed in Chapter 4; briefly, dogs are a naturally occurring stimulus class with individuals differing in multiple parts and relationships between parts, they share a first-order configuration, have a canonical upright, and information from shape-from-shading is potentially important.

Further, dog experts are a particularly good group of experts to test. Individual level discrimination is important when testing the expertise hypothesis (Diamond & Carey, 1986). Dog experts, especially judges and breeders, are used to looking at individual animals. Other classes of experts previously tested may not in fact do this. Car experts distinguish cars at the level of make/model and year (Gauthier et al., 2003). Biological cell experts also usually differentiate cells by type rather than as individuals. Bird experts usually make judgements at the level of species. When individual birds are identified it is usually on the basis of coloured bands on their legs, and/or behaviour, rather than the visual appearance of the bird itself (personal communications, from members of the Canberra Ornithological Group). This requirement of looking at the individual is also the main reason for not testing sheep experts (farmers in Australia typically have large flocks and are unlikely to know individual sheep), or flower experts.

Only a few classes of objects appear to be identified at an individual level by experts. These include dogs, cats and horses. Here I chose to test dogs. Given the results of Diamond and Carey (1986; Experiment 3), I matched the breed of dog to the breeds with which the experts have expertise. Labradors are a very popular breed in Australia and initial enquires suggested that this would be the breed for which most experts could be found (keeping in mind other constraints such as not using dogs with strong colour boundaries to parts). For these reasons, labradors and labrador experts were selected as best for the present experiments.

5.2.3.1 Subjects: characteristics of the labrador experts.

Fifteen experts participated. All reported normal or corrected-to-normal vision. Most experts were contacted from lists of qualified judges obtained from the Australian National Kennel Council website, or from lists of local breeders obtained from the ACT Canine Association website[2]. Additional recommendations were also obtained from some of the dog experts (most of those recommended were also on the ANKC list). All of the experts were Caucasian (the same race as the face pictures); seven were male. The experts ranged in age from 41-76 years, with most 55-66 years (mean age 58.2 years; median age 60 years). Highest education level was also recorded and ranged from early High school to postgraduate, with most having completed either Year 12 (High school) or tertiary education. Experts were tested in their own homes, except for two who preferred to be tested at the university, and were given a bottle of wine or book voucher to show appreciation for their time.

Level of expertise was assessed in a number of ways. All expert subjects except four were qualified All-Breeds judges (n=7) or Gun-dog Group[3] judges (n=4). Of the gun-dog judges, two also bred labradors, one had an animal transport business, and the fourth had handled labradors in the ring for friends. Of the four subjects who were not qualified judges, two bred labradors, one was involved in field-trials (in which labradors participate) and had owned labradors, and the fourth was a labrador guide-dog puppy trainer and dog-obedience trainer.

Subjects were asked several questions relating to experience with dogs in general, and labradors in particular. The aim of the questions was to ascertain the number of years subjects had been involved in various activities (e.g., judging, attending shows) and the number of individual labradors subjects would have seen in these activities. The method used was based on the personal history time-line method used to obtain long-term histories of drug use (Anglin, Hser, & Chou, 1993). This allows subjects to link levels of exposure with various important points in their life. In the drug use case, it has been shown to correlate fairly reliably with actual use. For most dog experts their involvement with dogs was fairly consistent over the years; however, the structure of the questions did seem to aid recall.

---

[2] http://www.ankc.aust.com/ and http://www.actca.asn.au/ respectively.

[3] Labradors are one of the four most common dogs in this group, the others being cocker spaniel, golden retriever, and Irish setter.

For each subject, the total number of years for which they had been involved with labradors and an estimate of the number of dogs seen was calculated (based on estimates of the number of years, shows per year and number of dogs per show, number of dogs bred etc.). Table 5.1 gives these estimates. Of these measures, the number of years is likely to be more reliable than the number of labradors seen. The full table including all estimates making up the total is given in Appendix III.

Table 5.1. Summary characteristics for each of the 15 dog experts (S1-S15) showing age, years of experience, dog shows attended per year and approximate number of labradors seen over the years.

	Age (years)	Years of experien ce	Dog shows /year	Approximate labradors seen
S1	76	37	25	11635
S2	66	25	20	1100
S3	46	42	50	10677
S4	59	30	12	3300
S5	64	32	12	3120
S6	55	23	6	2716
S7	56	23	6	2716
S8	60	34	25	8305
S9	60	10	45	400
S10	62	8	1	900
S11	70	8	1	900
S12	41	5	1	43
S13	51	19	20	2644
S14	63	22	6	4688
S15	44	28	20	15159
average	58.2	23.1	16.7	4553

From Table 5.1, the important points to note are that: all but three subjects had over 10 years' experience (exceptions had 5, 8, and 8 years); the overall mean was 23.1 years; and ten of the subjects had over 20 years' experience (mean 29.6 years for this sub-set). This compares favourably with previous studies of expertise (see Table 5.2 for the level of expertise from a range of studies). The dog experts tested here attended an average of 16.7 dog shows per year of which 4.9 were attended specifically with the purpose of looking at labradors (either to show or to judge). On average, experts had seen something in the order of 4550 labradors in their years involved with dogs. Note that the only expert who had seen less than 400 was the guide-dog trainer; although she had not seen as many dogs, she was personally familiar with those 43 individual dogs.

Table 5.2. Level of expertise reported in previous studies.

Study	Kind of expert	Average experience
Diamond & Carey (1986) Experiment 2	Dog	Not listed (listed as belonging to clubs etc.)
Experiment 3	Irish setters & cocker spaniel	Mean = 31 years
Rhodes & McLean (1990)	Birds	Not listed (mean rating as 5/7 on expertise scale)
Tanaka & Taylor (1991)	Bird	≥ 10 years, most > 20 years
	Dog	≥ 10 years, most > 20 years
Bruyer & Crispeels (1992)	Handwriting	Not listed (profession given as related)
Tanaka et al. (1996)	Biological cells	> 5-10 years
	Car	> 5-10 years
	Rottweiler dogs	> 5-10 years
Johnson & Mervis (1997)	Song birds	Not listed (mean peer rating of expertise >5/7)
	Tropical fish	Not listed (mean peer rating of expertise >5/7)
Gauthier et al. (2000)	Bird	Mean = 18 years
	Cars	Mean = 20.6 years
Tanaka & Curran (2001)	Bird	≥ 10 years, most > 20 years
	Dog	≥ 10 years, most > 20 years
Gauthier et al. (2003)	Cars	Not listed (expertise measured as the difference in sensitivity to matching cars and birds)
Gill-Spector et al. (2004)	Cars	As for Gauthier et al. (2003)
Rhodes, Byatt et al. (2004)	Lepidoptera (butterflies & moths)	3-30 years, mean 12.8 years

Anecdotal evidence of expertise was obtained during the experiment and in a post-test questionnaire. Throughout the experiments, experts were very good at picking which were American dogs, and several recognised that some of the Australian dogs were bred by a particular breeder. Many of the experts commented that even though there is a list of features (the breed standard) to which the dog must confirm, when they judge they look at the whole dog.

5.2.3.2 Structure of the testing sessions for experts.

Each expert participated in two sessions of approximately one hour each. Session 1 included an expertise questionnaire, the test of inversion on recognition memory (Experiment 7) and the contrast reversal test (Experiment 8). Session 2 consisted of the composite test (Experiment 9) and a post-test questionnaire to see whether experts personally recognised any of the dogs used in the experiments. Because experts participated in all three experiments, order of stimulus class (faces vs dogs), and orientation (upright vs inverted) was counterbalanced across subjects but remained the same for each subject across experiments. That is, if an expert saw upright dogs first in the memory test they also saw upright dogs first in the other two tests.[4]

5.3 Experiment 7 – Inversion Effects On Recognition Memory

Inversion effects are disproportionately large for faces compared to other objects when subjects have no specific expertise in that domain (Yin, 1969). In experts, Diamond and Carey (1986; Experiment 3) found inversion effects for dog experts looking at dogs that were as large as those for faces. This result has never been replicated and has not been found for other object classes (Bruyer & Crispeels, 1992; Rossion et al., 2002; Xu et al., 2004). Experiment 7 attempted to replicate Diamond and Carey's result for dog experts testing recognition memory in labrador experts as a function of stimulus class (labradors vs. faces) and orientation (upright vs. inverted).

A possible problem with Diamond and Carey's result is that dog pictures were taken from the archives of the American Kennel Club. Given that the experts were American Kennel Club judges it is possible that they were previously familiar with the

[4] The number of experts allowed one cycle of counterbalancing, plus one short of a second cycle. The final subject would have seen the dogs inverted first. A final subject was not tested because no further suitable (and available) experts could be found in the region.

actual photographs used and/or could individually name the dogs pictured. That is, the effect found may be due to specific pre-experimental familiarity with these dogs or pictures, rather than a general effect of expertise with the class.[5] To control for this possibility, in the current experiment, a post-test questionnaire was administered (after the composite task, Experiment 9). In this experts were presented with pictures of the dogs used in the experiments and asked to name each (or give relevant information such as breeder, or champion in a certain year). Of the 15 dog experts only 5 gave specific information about any dog. None correctly named any dog, although four each misnamed one dog as another which did appear in the experiments. Thus, any increase in inversion effects found in this experiment cannot be attributed to the experts having specific familiarity with the dogs tested.

5.3.1 Experiment 7 - Method

The method was the same as that used for novices (Chapter 4, Experiment 4) except that responses were recorded via a NewMicros Button Box connected to the iMac via a Keyspan USB twin serial adaptor. The yellow (middle) button was used for progressing through instructions and the buttons on left and right (red and green respectively) were used in the test phase to indicate which dog of a pair had appeared in the study phase.

5.3.2 Experiment 7 - Results

The percentage of test trials on which the "old" stimulus from an old-new pair was correctly chosen was calculated for each subject for each condition (2AFC, chance = 50%). These are shown averaged across expert subjects in Figure 5.1a. The results for novices from Chapter 4 are presented in Figure 5.1b for comparison. Note that novice subjects were mostly 18 years old whereas experts were mostly 55-66 years so overall

[5] To understand the difference, remember that disproportionate inversion effects for faces are found both for faces which are unknown prior to the experiment as well as for those which are familiar.

levels of accuracy are not comparable given that memory usually declines with age (e.g., Fastenau, Denburg, & Abeles, 1996).

Figure 5.2 contains RTs for experts and novices. These are the RTs for all trials including both correct and incorrect responses. (Results for correct-only responses were similar but, given the low accuracy in some conditions and the small number of trials, the RTs for correct-only responses were likely to be less reliable.)

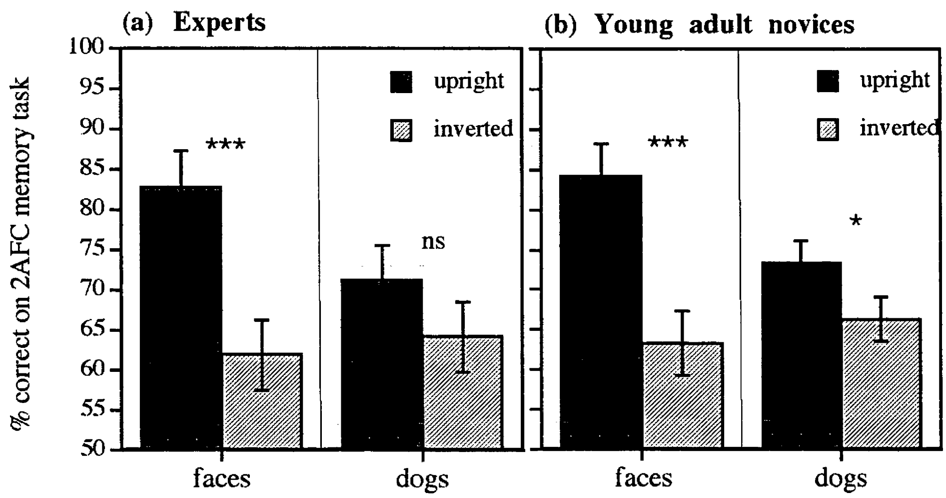


Figure 5.1. Experiment 8: Accuracy for the recognition memory test on faces and dogs, (a) expert subjects, (b) young adult novice subjects, N = 22, from Chapter 4 (Experiment 4). Error bars are appropriate for making the within subjects comparison between upright and inverted orientations (i.e., ± 1 SEM of the difference scores). *** $p < .001$, * $p < .05$, ns = $p > .05$.

5.3.2.1 Results for dog experts.

To show that the stimuli were well matched in conditions in which there should be equivalent expertise (i.e., essentially none) a first test compared faces and dogs in the inverted condition. As for novices, accuracy on inverted faces and inverted dogs did not significantly differ, $t < 1$.

A 2-way repeated-measures ANOVA for experts then revealed a significant main effect of orientation, $F(1, 14) = 15.41$, $MSE = 190.76$, $p < .01$, which was modified by a significant interaction between class (faces vs. dogs) and orientation (upright vs. inverted), $F(1, 14) = 6.95$, $MSE = 102.34$, $p < .05$. This reflects the fact that the inversion effect was much larger for faces (21 %) than for dogs (7 %; see Figure 5.1a). *A priori* t-tests showed that memory was significantly better in the upright than inverted orientation for faces, $t(14) = 4.69$, $p < .001$, but not for dogs, $t(14) = 1.62$, $p > .1$. In case a larger effect of inversion was being diluted by the fact that some experts had relatively few years of experience, the analysis was re-run with the sub-set of the experts who had over 20 years' experience (N=10). However, the inversion effect for dogs was still small (6 %) and non-significant, $t(9) = 1.30, p > .2$. Thus, dog experts did not show face-like inversion effects for dogs.

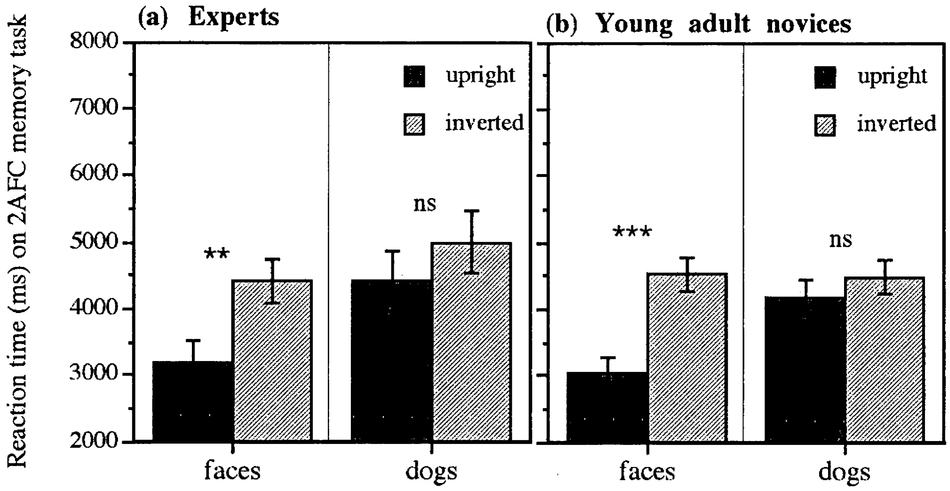


Figure 5.2. Experiment 8: Reaction times for the recognition memory test on faces and dogs, formatted as for Figure 5.1 (a) expert subjects, (b) young adult novice subjects, N = 22, from Chapter 4 (Experiment 4).

Results for RTs were also analysed. Analysis was conduced for all RTs and for correct RTs only. As the results were basically the same, and the number of correct trials in some cases approached chance, only results from analysing all RTs will be reported. A 2-way repeated measures ANOVA showed a significant main effect of orientation, $F(1, 14) = 8.68$, $MSE = 1404324$, $p <. 02$, as well as a significant main effect of class, $F(1, 14) = 8.69$, $MSE = 1413406$, $p <. 02$. The interaction between orientation and class was not significant, $F(1, 14) = 1.73$, $MSE = 909233$, $p > .2$, but *a priori* t-tests revealed that upright faces were recognised more quickly than inverted faces, $t(14) = 3.88, p < .01$, whereas there was no difference between upright and inverted dogs, $t(14) = 1.27, p > .2$. Given the relatively poor power associated with the RT data (RT is a very variable measure and there were only 15 trials for each of 15 experts), the RTs results were taken as sufficiently consistent with those for accuracy presented above.[6]

[6] Note that testing more trials to increase the reliability of RTs is not possible in a memory experiment as accuracy would be too close to floor. Note also that the number of trials tested here (15) is slightly more than used by Diamond & Carey (1986, Experiment 3) in which two lots of 6 trials were used.

5.3.2.2 <u>Comparison between dog experts and novices.</u>

The pattern of means for the experts was very similar to those of the novices in Chapter 4. On the accuracy measure, the statistical results were the same in both groups except that the small effect of inversion for dog-experts looking at dogs (7%, N=15) was not significant, whereas it had been for novices (7%, N=21). To compare experts directly to novices a 3-way mixed ANOVA was completed with expertise as the between subjects factor. The main effect of expertise was not significant, $F < 1$, and there were no interactions with expertise, all Fs < 1. A t-test comparing the size of the inversion effect for dogs (dogs upright – dogs inverted) for novices and experts also showed that there was no difference, $t < 1$. It is worth noting that the overall level of accuracy for dogs is similar for novices and experts (presumably because these are not age-matched controls), which has the advantage of guaranteeing a fair comparison between the size of the inversion effect for dogs.

Results from RTs (Figure 5.2) were again consistent with accuracy results. Experience did not interact with any other variables, Fs < 1, nor show any main effect, $F < 1$. Thus, dog experts did not have larger inversion effects in response to dogs than did novices.

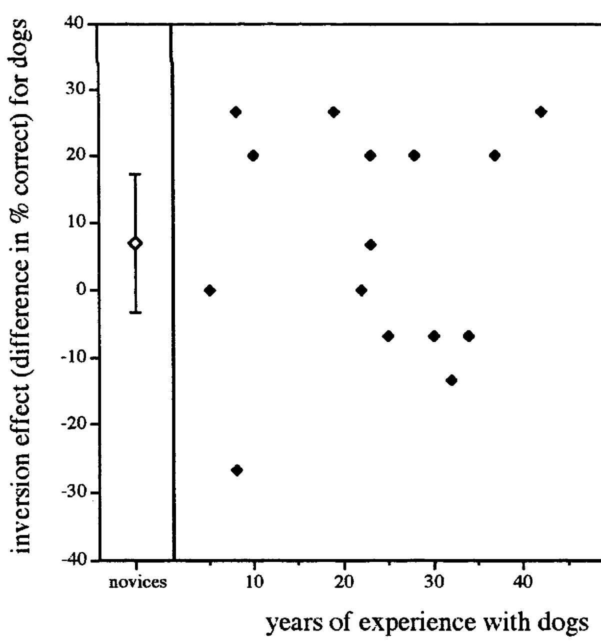


Figure 5.3. Experiment 7: Scatter plot of the size of the inversion effect (upright % correct - inverted % correct) for dogs versus the number of years experience. Each filled diamond represents one expert. The unfilled diamond on the left is the average for novice subjects ±1 SD.

5.3.2.3 Correlations with experience for experts.

To assess the effects of individual level of expertise, a scatter plot of expertise against the inversion effect for dogs was plotted (Figure 5.3). This showed that there was no relationship between the size of the inversion effect and years of experience, $r = .096, p > .5$. There was also no significant relationship between the size of the inversion effect and number of dogs seen (a much less reliable measure) although there was a small trend in this direction, $r = .324, p > .2$.

5.3.3 Experiment 7 - Discussion

For faces, dog experts showed the standard pattern of a large inversion effect on recognition memory. For dogs, however, experts showed the same small inversion effect as shown by novices (Chapter 4, Experiment 4), and not the large face-sized inversion effect reported by Diamond and Carey (1986, Experiment 3). The lack of any expertise effect was confirmed both by comparing the results of experts to novices, and also by testing for correlations between expertise and the size of the inversion effect for dogs.

Although the failure to replicate Experiment 3 of Diamond and Carey (1986) might initially seem surprising given that the result is often cited in the literature, it should be noted that Diamond and Carey's Experiment 2 did not find the face-sized inversion effect for dogs. Further, as reviewed in the introduction, face-like inversion effects for other objects-of-experts have not been found. I tentatively suggest that Diamond and Carey's finding of a face-sized inversion effect was due to subjects' specific familiarity with the dogs used as stimuli rather than a general effect of expertise.

There are, however, two possible concerns with the current results that deserve consideration. The first is a lack of overall difference in recognition performance between experts and novices on dogs, which could be taken to suggest that the experts were not really expert. However, given that memory is known to decrease with age, the fact that the older experts were as good as younger novices does not necessarily indicate a lack of expertise with dogs. Indeed, the lack of overall difference in performance between dog experts and novices may be due to experts having spared memory in both their areas of expertise (faces and dogs).

Data from age-matched controls[7] suggests that the age-based interpretation is correct. Control subjects were matched to individual experts on age (within 2 years), sex and education level. Results for the experts with more than 10 years experience (N=12) show the expected poorer performance for upright dogs in these age-matched novices compared to the experts (63% for novices vs. 74% for experts; $t(11) = 1.97, p < .04$, one-tailed), despite almost identical performance for faces (83% for novices vs. 82% for experts; $t < 1$). There will also be behavioural evidence that the dog experts are expert in Experiment 8.

The second possible concern is that, in Diamond and Carey's Experiment 3 experts' performance was matched for faces and dogs in experts in the upright orientation by using smaller learning sets for dogs. In the current experiment performance was matched for faces and dogs in the inverted orientation instead. This was done because it is the inverted condition for which experts should be similarly inexpert, whereas experience for upright faces is likely to exceed that for upright dogs even in dog experts. It could be suggested that the lack of a face-sized inversion effect was due this lack of matching performance in the upright orientation; however attempting to match upright performance in novices (the two additional experiments reported in Chapter 4) if anything decreased the size of the inversion effect (see Figure 4.3). Further, in the next experiment performance on upright dogs and faces for experts was matched, allowing this interpretation to be tested.

5.4 Experiment 8 – Contrast Reversal

Results from dog novices presented in Chapter 4 (Experiment 5) showed that contrast reversal had different effects in face and dog recognition. For faces, both contrast reversal and inversion were detrimental to performance (there were main effects of each) and these effects were independent of each other (there was no interaction).

---

[7] I designed this additional experiment; testing and some subject recruitment was conducted by a research assistant (Jacqui Brewer).

For dogs, the main effect of contrast reversal was much smaller, there was no main effect of inversion and no interaction. The aim of the present experiment was to test whether the results would be different for dog experts. Contrast reversal has previously only been tested with experiment-trained experts on artificial stimuli, and produced contradictory results: after training, greeble experts were significantly slower at recognising contrast reversed greebles than they had been before training (Gauthier et al., 1998), but blob experts showed no effect of contrast reversal on accuracy or RT (Nederhouser, et al., 2002).

The present experiment provides the first test of contrast reversal for real-world experts, and for a natural stimulus class. The task used was the same-different identity judgement of pairs of dogs or faces, as used for the novices in Chapter 4. Stimuli in each pair were both original contrast, both reversed contrast, or one original and one reversed contrast. If the effect of contrast reversal becomes face-like with expertise, then it is predicted that dog experts should show a difference in accuracy between original and reversed dogs that is larger than that found in novices, and potentially as large as the difference for experts between original and reversed faces.

5.4.1 Experiment 8 - Method

Methods were as for novices in Chapter 4 (Experiment 5), with one exception. With the young adult novices, the presentation time on each trial in Experiment 5 was 600 ms. Presentation time was increased for the older dog experts to keep performance off floor, and to achieve similar performance levels to the novices. In early testing, experts reported that they had trouble making fast enough eye movements to see both stimuli in a pair in the 600 ms allowed. For each expert, presentation time was adjusted in the practice phase until they felt that they could see both stimuli. The final time per trial ranged from 600-700 ms with longer durations generally given for older experts, and most tested at 650 ms.

5.4.2 Experiment 8 - Results

Table 5.3 presents mean accuracy in the same-different identity task for each condition[8]. As argued in Chapter 5, the both original and both reversed conditions are logically the most straightforward to interpret, and are plotted in Figure 5.4 (collapsed across same and different trials). In both Table 5.3 and Figure 5.4 the results from the young adult novices in Chapter 4 (Experiment 5) are included for comparison.

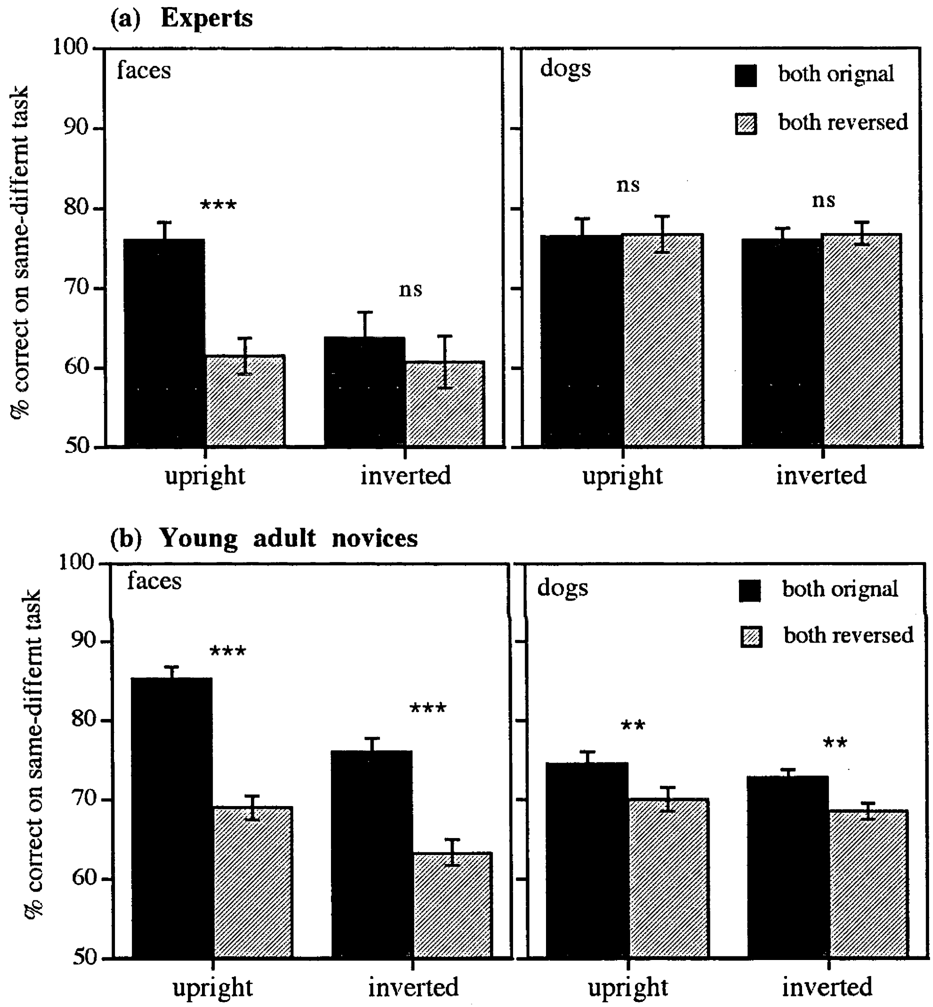


Figure 5.4. Experiment 8: Accuracy to say that both original contrast or both reversed contrast pairs are the same or different identity for faces and dogs for (a) experts and (b) novices from Chapter 4 (Experiment 5). Error bars are appropriate for the within subjects comparison of the two contrast conditions (i.e., ± 1 SEM of the difference scores). *** $p < .001$, * $p < .05$, ns = $p > .05$.

[8] There were no effects on RT. Given that the task involved limited presentation time, rather than a present-until-response procedure, this is not surprising or particularly interesting: under these conditions, reaction time is a largely meaningless measure.

5.4.2.1 Results for dog experts.

For experts, Figure 5.4 shows the usual substantial effect of contrast reversal for faces, but no contrast reversal effect for dogs in either orientation. A 2x2x3 repeated measures ANOVA involving stimulus class (dogs vs. faces), orientation (upright vs. inverted), and contrast condition (both original vs. both reversed vs. one original-one reversed) revealed a significant three-way interaction, $F(2, 28) = 3.57$, $MSE = 47.55$, $p < .05$. There was also a significant two-way interaction between stimulus class and contrast condition $F(2, 28) = 7.42$, $MSE = 44.21$, $p < .01$. Thus, further analysis was conducted separately for faces and dogs. The one original-one reversed condition was also excluded from further analysis, as the low-level differences between stimuli make it inherently harder than the other conditions (see Chapter 4).

Table 5.3. Mean accuracy (and SEM) in dog experts on the same-different task for each stimulus class, orientation, and contrast condition. Results from novices (Chapter 4, Experiment 5) are presented for comparison.

	Faces			Dogs		
	Both original	Both reversed	One original, One reversed	Both original	Both reversed	One original, One reversed
Dog Experts						
Upright	75.89 (3.79)	61.33 (2.96)	58.39 (1.78)	76.44 (2.97)	76.67 (3.10)	71.167 (2.13)
Inverted	63.56 (3.72)	60.67 (2.29)	58.39 (2.55)	76.00 (3.36)	76.67 (3.29)	71.00 (2.21)
Novices						
Upright	85.17 (1.58)	68.83 (2.19)	63.42 (1.46)	74.33 (2.36)	69.83 (1.84)	66.04 (1.39)
Inverted	76.00 (1.89)	63.25 (1.97)	59.79 (1.42)	72.67 (1.76)	68.42 (1.67)	65.04 (1.84)

For faces, a 2x2 ANOVA on the expert's data confirmed a main effect of contrast reversal, $F(1, 14) = 21.71$, $MSE = 52.57$, $p < .001$, and a main effect of orientation, $F(1, 14) = 9.84$, $MSE = 64.41$, $p < .01$. There was also a significant interaction, $F(1,14) = 7.89$, $MSE = 64.68$, $p < .02$. *A priori* t-tests showed that there were significant contrast reversal effects for upright faces, $t(14) = 6.70$, $p < .001$, but not for inverted faces, $t < 1$. *A priori* t-tests also showed that there was a significant inversion effect for the both original condition, $t(14) = 3.93$, $p < .01$, but not for the both reversed condition, $t < 1$. These results partially replicate the standard finding for upright faces (Novices in Chapter 4; Kemp, McManus, & Pigott, 1990), in that faces show a large contrast reversal effect. One difference from the previous results was the smaller contrast effect for inverted faces than for upright (i.e., a significant interaction). While this could be meaningful, it was not previously obtained with the dog novices where performance was better, and could have arisen here simply because performance for inverted faces is approaching floor.

For dogs, 2x2 ANOVA on the expert's data revealed no main effects or interaction, all Fs < 1. *A priori* t-tests showed that there was no effect of contrast reversal in either orientation, ts < 1, or of orientation for either of the contrast conditions, ts < 1. The results for the sub-set of experts with over 20 years' experience (N=10) were basically the same. In particular, t-tests again showed no effect of contrast in either orientation, ts < 1, or of orientation for either contrast condition, ts < 1. Thus, dog experts did not show the same pattern for dogs as for faces, with respect to either contrast reversal or inversion.

5.4.2.2 Comparison between dog experts and novices.

There was no indication that dog experts looking at dogs were more strongly affected by either contrast reversal or inversion than novices (Figure 5.4). This was confirmed by a 4-way mixed ANOVA directly comparing experts and novices as the between subjects factor, and including class (faces vs. dogs), orientation (upright vs. inverted), and contrast condition (both original vs. both reversed) as within subjects factors. The ANOVA showed no 4-way interaction involving experience, $F(1,33) = 1.87$, $p > .18$, and the relevant 3-way interactions involving experience were also far from significant. In particular, the pattern of 2-way interaction between class and contrast condition (i.e., the larger contrast reversal effect for faces than dogs) did not differ between experts and novices, as demonstrated by the lack of class x contrast x experience interaction, $F < 1$; and the pattern of 2-way interaction between class and orientation condition (i.e., the larger inversion effect for faces than dogs) did not differ between experts and novices, as demonstrated by the lack of class x orientation x experience interaction, $F < 1$. Thus, there was no evidence that dog experts developed contrast reversal or inversion effects for dogs that were any different from those found in novices.

The 4-way ANOVA did reveal two significant interactions involving expertise. First, collapsing over both contrast condition and orientation, there was a 2-way interaction between class and experience, $F(1, 33) = 24.08, p < .001$. This is important in that it provides behavioural evidence of expertise in the dog experts. Recall that the overall levels of performance are not directly comparable, because the dog experts were older and also given more time. However, despite the fact that the (older) dog experts were significantly poorer than the (younger) novices with faces (experts = 63 %, novices = 69 %; $t(33) = 2.40, p < .05$), the experts tended to be better than the novices with dogs (experts = 75 %, novices = 69 %; $t(33) = 1.93, p = .06$). This pattern was also confirmed when the upright both original contrast condition was considered alone, namely the condition for which dog experts should have most expertise. There was again a class x experience interaction, $F(1, 33) = 6.89, p < .02$, reflecting poorer performance in experts than novices for faces (experts = 76 %, novices = 85 %, $t(18.88) = 2.26, p < .05$, with degrees of freedom adjusted for unequal variance), but somewhat better performance in experts than novices for dogs (experts = 73 %, novices = 69 %, $t < 1$).

The second interaction from the 4-way ANOVA was that, collapsing over class and orientation, there was a 2-way contrast x experience interaction, $F(1, 33) = 22.22$, $p < .001$. As can be seen in Figure 5.4, this reflects the fact that experts showed slightly smaller contrast reversal effects than novices, for both faces and dogs, and in both orientation conditions. This result is of little interest in itself. More important is the finding that when upright dogs are considered alone, the contrast reversal effect (both original – both reversed) did not differ between experts and novices, $t(33) = 1.81$, $p > .07$, and indeed the non-significant trend was towards a smaller contrast reversal effect for dog experts. The result was the same when experts with more than 20 years' experience were considered, with novices again tending to show a larger contrast reversal effect, $t(28) = 2.02, p = .053$.

5.4.2.3 Correlations with experience for experts.

Within the expert group, a scatter-plot of expertise against the contrast reversal effect (see Figure 5.5) showed no relationship between size of the effect for upright dogs and years of experience, $r = -.16$, $p > .5$. There was also no relationship with number of dogs seen, $r = -.10$, $p > .7$.

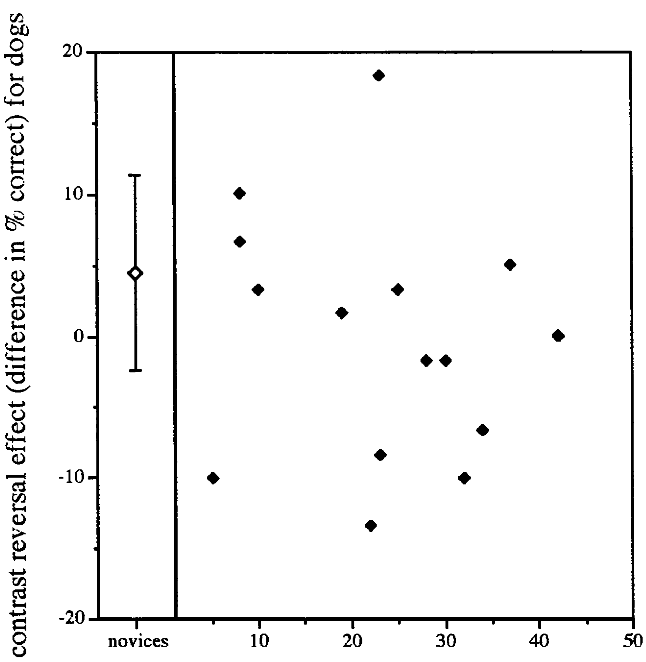


Figure 5.5. Experiment 8: Scatter plot of the size of the contrast reversal effect (both original % correct - both reversed % correct) for dogs versus the number of years experience. Formatting as in Figure 5.3.

5.4.3 Experiment 8 - Discussion

In a same-different identity judgement task, dog experts were affected by both contrast reversal and inversion for faces, replicating previous results (Chapter 4; Kemp et al., 1990). For dogs, however, experts showed no effect of either contrast reversal or orientation. Further, amount of experience was not related to the size of the contrast reversal effect, and results for experts and novices did not differ in any way that suggested larger contrast reversal or inversion effects in experts.

Experts and novices differed in only two ways. Experts were slightly less sensitive to contrast reversal than novices and this applied to all stimuli (faces, dogs, upright, inverted), rather than showing a greater contrast effect restricted to upright dogs as would be predicted by the expertise hypothesis. Secondly, dog experts performed relatively better with dogs than with faces as compared to novices, confirming their expertise via behavioural performance.

In conclusion, results of Experiment 8 show that, while experts had enhanced overall performance with their objects-of-expertise as compared to novices, standard patterns of "face-specific" processing did not emerge for labradors. Of the studies to previously test contrast reversal with expertise, only Gauthier et al. (1998) found an effect with greeble experts. Nederhouser, et al. (2002), conversely, found no effect of contrast with blob experts. In the present result, dog experts were if anything slightly less affected by contrast reversal than novices, even though the stimuli were natural objects, and of a class where shape-from-shading is potentially useful.

As a final point, inversion effects were also tested in this experiment. In the present experiment, unlike in Experiment 7, performance for faces and dogs was equivalent in dog experts in the upright orientation (see white bars in Figure 5.4). Despite this fact, face-size inversion effects for dogs were not found. This argues that a failure to equate performance for upright faces and dogs cannot explain the results of Experiment 7.

5.5 Experiment 9 – Composite Effect

Experiment 9 was a direct test of holistic processing using the composite test (Young et al., 1987). The task was simultaneous matching, as suitable for unfamiliar faces and dogs, and results from aligned and unaligned composites were contrasted (i.e., the same task and comparisons as for the novices tested in Chapter 4, Experiment 6). Composites were made so that the half-to-compare (e.g., forehead) could be the same or different, but the half-to-ignore (e.g., chin) was always different. For the novices, in Chapter 4, for upright faces it was harder to correctly say that the half-to-compare was the <u>same</u> if the two halves were aligned than if they were unaligned. For inverted faces, and for dogs (both upright and inverted) there was no difference between aligned and unaligned conditions. That is, the composite effect (i.e., holistic processing) occurred only for upright faces, and not for dogs or inverted faces.

For experts, the composite test has only previously been used with greebles. No reliable evidence of a composite effect was found (Gauthier & Tarr, 2002; Gauthier et al., 1998). This is the first test of holistic processing using the composite paradigm in real-world experts. For face stimuli, results are predicted to be the same for dog experts as for novices. That is, accuracy should be lower for aligned than unaligned trials but only in the upright orientation. The question of interest is then whether dog experts show a composite effect for dogs. A lack of composite effect, and thus the same results as for novices, would indicate no holistic processing for objects-of-expertise. In contrast, a composite effect (aligned harder than unaligned) for dogs, and a corresponding interaction with level of experience, would indicate that objects-of-expertise are holistically processed. As in the previous chapter, the focus will be on the "same" trials where the predictions are clear.

5.5.1 Experiment 9 - Method

The method was the same as for novices (Chapter 4, Experiment 6) with the following changes. Expert subjects were given extra time to view the stimuli, adjusted for each subject as in Experiment 8 (contrast reversal test). Times per trial ranged from 650-800 ms and, as with Experiment 8, longer times were used for older experts. After testing 12 experts it was noticed that some had a strong bias to respond same rather than different, particularly for dogs, meaning that scores in some of the conditions of interest were approaching ceiling. Two methods to overcome this were tried. The first was to give subjects longer to view the stimuli. This was used with one subject (850 ms) but did not remove bias. A version of the experiment was then created which included feedback (a beep on incorrect responses) to discourage bias. This was pilot tested with five novice subjects similar in characteristics to those described in Chapter 4 (at 600 ms to match the original novices). These five showed less bias than the original novices. Thus, the last two experts were tested with auditory feedback. As the pilot subjects also showed general improvement, time for the last two experts was reduced again to 650 ms and 600 ms respectively. Note that the primary comparison is between aligned and unaligned conditions which is fully within subjects. Thus, these procedural differences between subjects cannot affect the presence or otherwise of a composite effect.

5.5.2 Experiment 9 - Results

Accuracy in matching halves (percentage of "same" responses for same-identity trials, and percentage of "different" responses for different-identity trials) was calculated for each subject for each condition (results are presented in Table 5.4 separated by half-to-match). A 4-way repeated measures ANOVA showed that patterns differed across same and different trials with a significant 4-way interaction between stimulus class, orientation, alignment, and trial type, $F(1, 14) = 5.23$, $MSE = 26.27$, $p < .05$. From a theoretical perspective, there are clear predictions corresponding to a composite effect for same trials, but not for different trials (see Chapter 4); Figure 5.6 gives the results for same trials for both faces and dogs. Results of the two sets of young adult novices tested in Chapter 4 (Experiment 6a & 6b) are provided for comparison.

5.5.2.1 Results for dog experts.

For the dog experts (Figure 5.6a), a 3-way repeated measures ANOVA (class x alignment x orientation for same trials) revealed a non-significant 3-way interaction, $F(1, 14) = 1.75$, $MSE = 28.04$, $p > .2$, but a significant 2-way interaction between orientation and alignment, $F(1, 14) = 12.43$, $MSE = 11.33$, $p < .01$, and a 2-way interaction between class and alignment that approached significance, $F(1, 14) = 3.78$, $MSE = 15.93$, $p = .072$. Figure 5.6a indicates that these results reflect a pattern in which there was a composite effect only for upright faces. To assess possible composite effects for each stimulus class directly, dogs and faces were analysed separately. For faces, a 2-way ANOVA confirmed a significant interaction between alignment and orientation, $F(1, 14) = 7.54$, $MSE = 23.60$, $p < .02$. *A priori* t-tests showed that accuracy was significantly lower for aligned trials than unaligned trials for upright faces, $t(14) = 3.24$, $p < .01$, but not for inverted faces, $t < 1$. Thus, as expected, there was a significant composite effect for upright faces but not inverted faces.

For dogs, there was no orientation x alignment interaction, $F < 1$, or main effects of either alignment or orientation, Fs < 1. *A priori* t-tests also showed no differences between aligned and unaligned trials, either for upright dogs, $t < 1$, or inverted dogs $t(14) = 1.44$, $p > .15$ (the small trend was in the reverse direction for a composite effect). Thus, dog experts showed no evidence of a composite effect for dogs in either orientation.

Table 5.4. Experiment 9: Mean accuracy (SEM) in dog experts for the composite task shown separately for (a) same and (b) different trials, and within each, by class, orientation, half-to-compare, and alignment. Results from Experiment 6a & 6b are also shown.

a) Same trials

	Faces				Dogs			
	Top half		Bottom half		Top half		Bottom half	
	Aligned	Unalign	Aligned	Unalign	Aligned	Unalign	Aligned	Unalign
Dog Experts								
Upright	79.1	82.1	79.6	88.8	90.0	90.2	85.6	86.8
	(3.1)	(3.9)	(3.4)	(2.0)	(2.6)	(1.8)	(2.4)	(2.3)
Inverted	86.8	86.7	85.2	83.8	92.4	90.6	84.7	84.4
	(2.9)	(2.9)	(2.4)	(2.9)	(1.9)	(2.2)	(3.5)	(4.3)
Novices (6b)								
Upright	80.6	85.3	82.8	86.4	90.2	87.8	84.1	84.8
	(2.9)	(2.7)	(3.2)	(3.3)	(2.4)	(2.9)	(3.7)	(4.2)
Inverted	85.9	84.1	81.2	83.6	90.7	90.8	86.6	84.8
	(2.9)	(3.6)	(3.9)	(3.8)	(1.7)	(1.6)	(3.2)	(3.8)
Novices (6a)								
Upright	-	-	-	-	87.1	88.6	83.8	83.9
					(1.8)	(1.5)	(2.3)	(2.4)
Inverted	-	-	-	-	84.5	84.5	79.5	78.6
					(2.3)	(2.7)	(2.7)	(2.8)

b) Different trials

	Faces				Dogs			
	Top half		Bottom half		Top half		Bottom half	
	Aligned	Unalign	Aligned	Unalign	Aligned	Unalign	Aligned	Unalign
Dog Experts								
Upright	69.8	66.9	69.1	61.6	71.3	74.0	44.4	46.7
	(6.0)	(5.7)	(5.5)	(4.8)	(4.0)	(4.0)	(6.1)	(6.0)
Inverted	59.1	55.4	56.9	56.4	63.6	59.9	42.4	43.3
	(6.6)	(6.4)	(6.3)	(6.5)	(5.5)	(5.8)	(6.2)	(5.9)
Novices (6b)								
Upright	80.2	75.6	69.6	63.9	67.2	71.2	55.0	58.0
	(3.5)	(3.2)	(5.1)	(5.3)	(4.2)	(4.1)	(5.2)	(5.2)
Inverted	68.0	66.7	58.0	55.3	58.6	60.5	55.2	54.5
	(3.6)	(3.9)	(5.1)	(5.2)	(5.5)	(5.5)	(5.4)	(5.3)
Novices (6a)								
Upright	-	-	-	-	74.2	74.0	64.6	65.6
					(2.9)	(3.0)	(4.4)	(4.2)
Inverted	-	-	-	-	67.8	69.0	64.4	61.9
					(2.9)	(3.2)	(3.5)	(3.9)

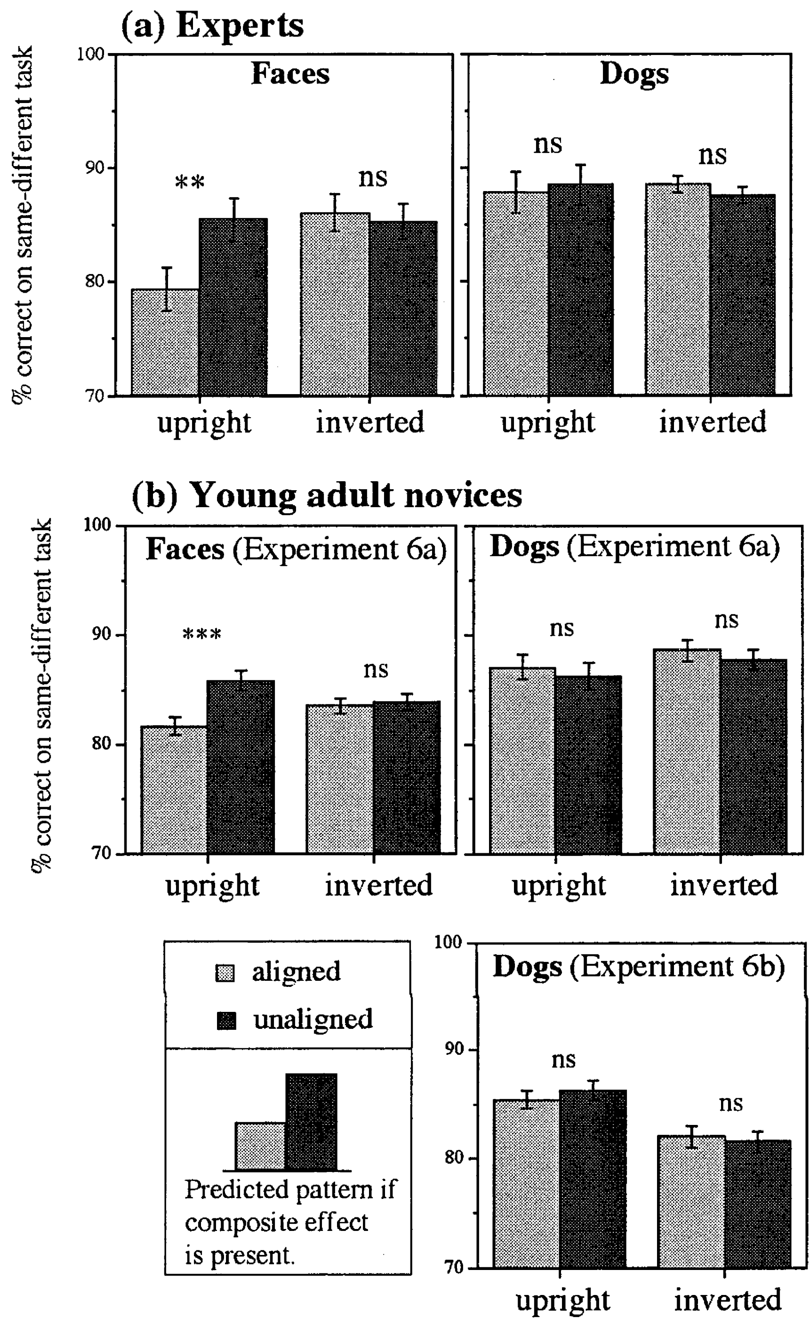


Figure 5.6. Experiment 9: Accuracy to compare target-half of a pair of composites (simultaneous presentation same-different task). Results are for "same" trials only, collapsed across top-half-to-compare and bottom-half-to-compare trials. These are shown for (a) experts and (b) both groups of novices from Chapter 4 (Experiment 6). Error bars are appropriate for the within subjects comparison of aligned and unaligned conditions (i.e., ± 1 SEM of the difference scores). *** $p < .001$, * $p < .05$, ns = $p > .05$. The predicted pattern for a composite effect, indicating holistic processing, is that accuracy should be lower for aligned than unaligned trials.

These analyses were conducted collapsing over top-half-to-compare and bottom-half-to-compare to maximise power. Results for top and bottom halves were also analysed separately. For faces, the trends described above were present for both halves with the only substantive difference being that the composite effect for upright faces was significant only for bottom halves (chins), $t(14) = 3.21$, $p < .01$, and was in the predicted direction but did not reach significance for top halves, $t(14) = 1.69$, $p > .1$. For dogs there was no composite effect, either for top-half-to-compare ($t < 1$ upright; $t(14) = 1.97$, $p > .06$ inverted in the reverse direction to a composite effect) or bottom-half-to-compare ($t < 1$ upright; $t < 1$ inverted). Thus, it was not the case that a composite effect for dogs had been hidden by collapsing over half-to-compare.

5.5.2.2 Were experts too close to ceiling?

A possible concern with the lack of a composite effect for dogs was that experts' mean performance on same trials was closer to ceiling than might be desired (Figure 5.6). To demonstrate that the lack of a difference between aligned and unaligned trials was not occurring for subjects further from ceiling, a sub-set of experts whose accuracy was 90% or less on the unaligned upright dogs condition were chosen for analysis (see Figure 5.7). This criterion was chosen because (a) if holistic processing were to occur for dogs accuracy would be higher for unaligned trials than aligned trials (thus shifting any effect further away from ceiling); and (b) theoretically, holistic processing would occur for upright dogs if anywhere as this is the orientation in which dog experts usually experience dogs. Data were collapsed over top and bottom halves, with mean performance for upright unaligned dogs (84.7%) now comfortably below ceiling (indeed it was slightly lower than the 85.4% for faces in the original all-experts analysis). As shown in Figure 5.7, Results did not differ in any important way from those reported above. Most importantly the *a priori* t-test comparing unaligned and aligned trials for upright dogs was not significant, $t < 1$, while the effect for upright faces was still present, $t(9) = 2.49$, $p < .05$. Thus, the finding of no composite effect for dog experts looking at dogs cannot be attributed to ceiling effects.

5.5.2.3 Comparison between dog experts and novices.

Experts were compared to each of the two groups of novices from Chapter 4 (Experiment 6) in turn. Comparisons were only made for dogs because of large differences in variance between the novices and experts for faces. There was no suggestion that composite effects for dogs were any larger in experts than in novices. For dogs, there were no main effects or interactions with experience for either group, all

$ps > .11$. *A priori* t-tests comparing the size of the composite effect (unaligned - aligned) for upright dogs showed no difference between experts and either group of novices, $t < 1$ in both cases. It is worth noting that, as intended, overall accuracy on dogs for experts and novices was very similar (collapsed over orientation: experts = 88.%; novices group 1 = 84%; novices group 2 = 88%)[9].

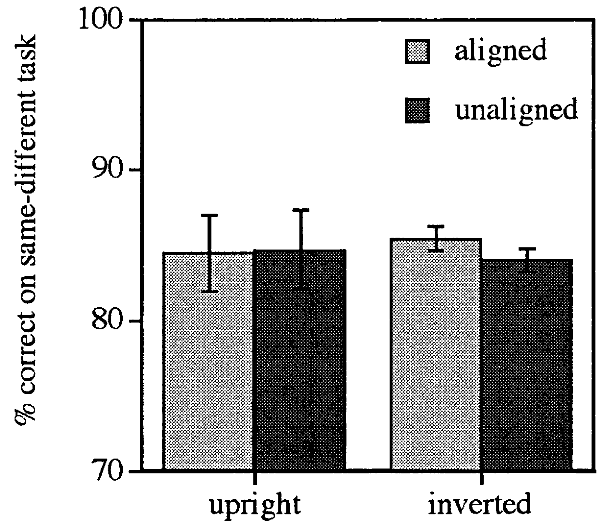


Figure 5.7. Experiment 9: Results for the aligned and unaligned composite dogs (collapsed over half-to-compare) for a sub-set of experts (n=10) experts whose accuracy was 90% or less on the unaligned upright dogs condition. Formatting as for Figure 5.6.

5.5.2.4 Correlations with experience for experts.

A scatter plot for the size of the composite effect for upright dogs against the number of years' experience is presented in Figure 5.8. This did not show any increase in the composite effect with increasing experience. Indeed, the non-significant trend was in the opposite direction, $r = -.48, p = .069$. This was also the case for the size of the composite effect compared to number of dogs seen, $r = -.48, p = .070$.

---

[9] Behavioural effects of expertise as found in Experiment 8 would not be expected here given that the task involved decisions about half dogs.

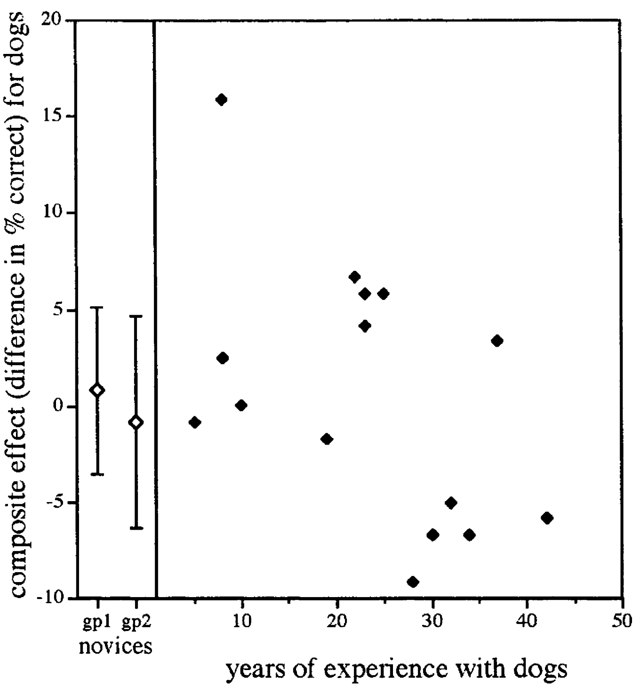


Figure 5.8. Experiment 9: Scatter plot of the size of the composite effect (unaligned correct % - aligned % correct) for dogs versus the number of years experience. Each filled diamond represents one expert. The unfilled diamonds on the left represent averages for novice subjects (±1 SD) for each of the two groups tested (gp1 = subjects in Experiment 6a, while gp2 = subjects in Experiment 6b).

5.5.3 Experiment 9 - Discussion

As expected, Experiment 9 revealed a composite effect for upright faces, and no composite effect for inverted faces, replicating previous findings (novices in Chapter 4; Carey & Diamond, 1994; Le Grand, et al., 2004; Young et al., 1987). For dogs, in contrast, dog experts showed no sign of a composite effect even when the dogs were upright. There was also no difference between novices and experts, and if anything there was a slight negative correlation between level of expertise and the size of the composite effect for dogs. The composite effect is a particularly strong direct test of holistic processing (see Chapter 4). Thus, Experiment 9 shows that expertise does not produce holistic processing of labradors.

5.6 General Discussion

The aim of the present studies was to assess whether many years of practice at discriminating individual labradors would lead to face-like processing of those labradors. The results of three experiments argue strongly that this does not occur, even with an average of 23 years' experience. In a test of inversion effects on recognition memory, dog experts showed much smaller inversion effects for dogs than for faces, and their results did not differ from those of novices. In a test of the effects of contrast reversal, dog experts were affected much less by contrast reversal of dogs than of faces. Further, they showed, if anything, less of an effect of contrast reversal on dogs than did novices. Finally, in a direct test of holistic processing, using Young et al's (1987) composite effect, dog experts holistically processed upright faces, but not inverted faces or dogs in either orientation. Results for dog experts looking at dogs again did not differ from those of novices. Further, in no case did the measure of face-like processing correlate with experience; a small but non-significant trend was found for the inversion effect (as might be expected) but there was none whatsoever for contrast reversal or the composite effect. The current studies provide strong evidence against the expertise hypothesis; that is, against the claim that extensive experience at making individual discriminations for any object class produces face-like processing.

5.6.1 Nature of the experts tested.

The present results are particularly strong because of the experts used. As noted in the introduction, not all experts make individual level discriminations on the basis of visual characteristics. Dog experts do (as do cat experts), whereas bird and flower experts, for example, do not. Further, the difference between Diamond and Carey's (1986) Experiment 2, in which they did not find a face-sized inversion effect, and Experiment 3, in which they did, had suggested that experts must be expert with the particular sub-type (e.g., breed of dog) being used as stimuli. There is evidence that the current experts fulfilled this criterion. Empirical evidence came from Experiment 6, where experts were better with dogs than age-matched novices, and Experiment 7, where experts were relatively better with dogs than novices, despite being worse with faces. There was also anecdotal evidence from the fact that experts could recognise dogs as American versus Australian and as bred by a particular breeder. Dog experts are also a good choice of expert because dogs make a good comparison stimulus for faces: dogs are natural stimuli and, due to genetic variation, differences between individuals can occur in multiple locations all over the dog. This is more likely to encourage holistic processing than comparisons between individuals which only differ in terms of a few specific parts (e.g., greebles).

In terms of years of experience, there is also good reason to think that the experts in the present experiments had sufficient expertise to provide a fair test of the expertise hypothesis. The experts in these studies, as well as having more experience than those in many previous studies (see Table 5.2), had on average approximately the same number of years looking at dogs (mean experience = 23 years, mean age = 58 years) as the younger novices (mean age = 22 years) had looking at faces. This may not, of course, mean that the dog experts had seen as many dogs as the dog novices had seen faces. However, it seems likely that the dog experts would have seen at least as many dogs as a school age child had seen faces. Literature on the development of face processing has shown holistic processing in 6-year-olds (composite effect, Carey & Diamond, 1994; part-whole effect, Tanaka et al., 1998; second order-relational processing, Gilchrist & McKone, 2003; Mondloch et al., 2002) and even 4-year-olds (part-whole effect, Pelicano & Rhodes 2003; second order-relational processing, McKone & Boyer, submitted). Thus, the present experts should easily have had sufficient expertise to test the hypothesis that expertise with a non-face object class will cause face-like processing to develop.

5.6.2 Comparison of the present results with previous studies.

The current finding, that expertise did not lead to a larger inversion effect for objects-of-expertise (in this case dogs) than that found in novices, is consistent with most of the previous literature (Gauthier et al, 1998; Gauthier, Skudlarski et al., 2000; Xu et al., 2004, when % correct is used as the measure). The inversion effect for dogs in experts was also much smaller than that for faces (consistent with references in the previous sentence as well as Bruyer & Crispeels, 1992, and Rossion et al., 2002, where small increases with expertise were found).

The only previous study that the current results are inconsistent with is Experiment 3 of Diamond and Carey (1986), where dog experts showed the same sized inversion effects for faces and dogs (of their breed of expertise). As noted above, the experts in the current experiment were experts with the breed tested, but the same-sized effect was not found for faces and dogs even when a sub-group of experts with the same number of years' experience as Diamond and Carey's was used (30 years in the present study vs. 31 in theirs), or when performance for upright faces and upright dogs was matched (Experiment 8) as theirs was. So why was the face-sized inversion effect for dogs not replicated? As mentioned in the introduction and discussion of Experiment 7, the most likely explanation is that Diamond and Carey's experts knew the individual dogs used as stimuli, and possibly even particular images of those dogs. The pictures used in Diamond and Carey's study were acquired from the American Kennel Club, which was also the organization through which dog experts were contacted. It is thus entirely possible that those images formed part of the set used by the American Kennel Club to train judges. Being able to name images (or dogs) based on pre-experimental knowledge would increase memory performance but this would, at least initially, be specific to the orientation in which the images were learnt (cf., naming rotated object studies, for example, Jolicoeur, 1985). In the present study, in contrast, experts' familiarity with particular dogs/images was tested, and it was ensured that none of the dog experts could correctly name any of the dogs. Overall, I suggest that the single finding of a face-size inversion effect for experts viewing objects-of-expertise can be better explained by specific familiarity than by generic expertise. Results from this paradigm do not, therefore, generally support the expertise hypothesis.

With respect to the composite task, the current studies provide the first test of this paradigm with a natural object class and the first test with real-world experts. The finding of no composite effect for dogs in dog experts is entirely consistent with the only previous studies of the composite effect in experts; that is, with the finding of no composite effect for greebles in greeble experts (Gauthier & Tarr, 2002; Gauthier et al., 1998). The composite effect provides a stronger test of the expertise hypothesis than the size of the inversion effect, because the composite test directly assesses holistic processing. Further, as argued in Chapter 4, the composite effect is a particularly strong test of holistic processing as there is no composite effect for objects in non-experts rather than simply a smaller effect than occurs for faces (as in the part-whole task). Thus, the present result of no composite effect for dog experts looking at dogs, combined with the previous same result for greebles, suggests very strongly that generic expertise does not lead to holistic processing.

The effects of contrast reversal on objects-of-expertise have also not previously been tested for any natural stimulus class or real-world experts. The current finding of no effect of contrast reversal on objects-of-expertise is consistent with a previous finding of no effect of contrast reversal for blob experts (Nederhouser et al., 2002) but not with a finding of an effect for greeble experts (Gauthier et al., 1998). One difference between the studies is that both Nederhouser et al. and the present study used a matching task, whereas Gauthier et al. used a naming paradigm. However, given that matching produces a clear contrast reversal effect for faces, there seems no reason why it should not produce a contrast reversal effect for objects-of-expertise if one were there to be found.

Results from the present study are thus generally consistent with results from previous studies. This is true for all three paradigms tested.

5.6.3 Overall status of the expertise hypothesis.

The present results, as well as those of previous research strongly suggest that face-like processing, as tested on a range of behavioural tasks, is not simply a matter of extensive experience. In terms of previous behavioural studies the results are also consistent with findings from the part-whole effect. As reviewed in the introduction, part-whole effects are no larger for experts looking at their object-of-expertise than they are for novices. This has been tested for a range of manmade (car fronts, greebles) and natural objects (dog faces, biological cells). Of course, some aspects of cognitive processing for objects do change with expertise. For example, Rhodes and McLean (1990) showed that for bird experts, caricatures of very similar birds were recognised more quickly than "anti-caricatures", but this difference was not apparent for novices (looking at more dissimilar birds). Similarly, experts show a downward shift in their level of classification compared to non-experts; that is, they are more likely to classify objects-of-expertise at a subordinate level rather than at the basic level (K. E. Johnson & Mervis, 1997; Tanaka & Taylor, 1991). However, the styles of processing which are special to faces in nonexperts – holistic processing and particular sensitivity to shape from shading information – do not seem to occur for objects-of-expertise.

There is also little support for the expertise hypothesis from neuropsychology and neuroimaging (fMRI, ERP and MEG). Neuropsychology suggests a double-dissociation between the processing of faces and objects of expertise (Dixon et al., 2002; Duchaine et al., 2004; McNeil & Warrington, 1993; Moscovitch et al., 1997). Similarly, neuroimaging suggests that the area of the brain used most specifically for face processing (the FFA) is still substantially less activated for objects-of-expertise than for faces (Grill-Spector et al., 2004) and that very few of the voxels in the area which are activated by objects-of-expertise are also activated by faces (Rhodes, Byatt et al., 2004). Similarly, although there is a strong N170 response to some objects-of-expertise it is not in the same area of the brain as that found for faces (e.g., Rossion et al, 2002; Tanaka & Curran, 2001).

The conclusion that the available data does not support the expertise hypothesis may seem surprising to some readers, given that the current "zeitgeist" is in favour of the expertise hypothesis. I see three general sources of confusion in this area. The first is that a number of critical results have been widely cited as supporting the expertise hypothesis when in fact their original results were much weaker than implied. In an extreme example, Elgar and Campbell (2001, p.29), discussing the face-specific of areas of the inferotemporal cortex, claim that Gauthier et al. (1999) have "shown that expertise in the (sub-categorical) identification of *any* visual material makes use of this substrate" (their emphasis). The study in question tested only one object class (greebles), and as discussed in the introduction more recent studies (Grill-Spector et al., 2004; Rhodes, Byatt et al., 2004) have brought the results (and those of Gauthier, Skudlarski et al., 2000) into question. More moderately, Burgund and Marsolek (2000, note 4, p.489) cite Gauthier and Tarr (1997) as having shown "greater whole-based processing of the stimuli" with expertise; however, as detailed in the introduction to this chapter, part-whole effects for experts in Gauthier and Tarr (1997; or Gauthier & Tarr, 2002; Gauthier et al., 1998) were not reliably larger than in novices. Le Grand, et al. (2004, p.768) similarly cite Gauthier and Tarr (2002) as having shown that greeble experts "show evidence of holistic processing in the composite task". Again, as discussed in the introduction, there is no evidence of a composite effect in greeble experts.

Some studies have also been purely miscited. For example, Rouse, Donnelly, Hadwin, and Brown (2004, p.1) cite Gauthier, Skudlarski et al (2000) as evidence that "development brings an increasing ability to process faces configurally rather than in a piecemeal fashion". Gauthier, Skudlarski et al (2000) studied whether activation in the FFA was higher for objects-of-expertise in bird and car experts than in novices, and did not directly measure either configural/holistic processing or development. Watanabe, Kakigi, & Puce (2003, p.879) cite Valentine (1988) as evidence that "the face inversion effect has been attributed to our 'expertise' with a highly homogeneous class of stimulus seen in one orientation in everyday life". In fact, Valentine noted that the evidence available to him was inconclusive. Haggbloom & Warnick (2003, p.579) cite Gauthier and Tarr (1997) as one source of evidence for the statement that "an acquired expertise for the recognition of non-face stimuli produces a large inversion effect". The sentence seems to suggest that inversion effects increase with expertise. I have no problem with the statement in general, but Gauthier and Tarr (1997) did not test inversion effects (they tested the part-whole effect in both orientations).

The second source of confusion is that, as noted above, there are some aspects of object processing which do change with experience, but these are not the aspects which are special to face processing. Mentioned above were caricature effects, and a downward shift in level of classification. Another aspect of object processing that changes with extensive practice is inversion effects. For objects these generally disappear rapidly with practice (for review, see McKone & Grenfell, 1999). Conversely, inverted faces are still processed differently from upright faces even after extensive practice (as shown in Chapter 3).

Finally, there is also confusion about what "holistic' processing means. An important example of this is the paper by Gauthier et al. (2003) in which holistic processing is defined as "obligatory processing of all features of an object" (p.428). This paper has been cited as showing a composite effect for objects-of-expertise. Le Grand et al (2004) cite Gauthier et al. (2003) as having found "behavioural interference of car processing on holistic face processing, as indicated by an attenuated composite face effect". In their paper Gauthier et al. presented car "composites" made by combining the top half of one car with the bottom half of another. These were always presented aligned, but either with the top half upright or inverted. The task was a sequential same/different matching task (on the bottom half). Gauthier et al. measured "holistic" processing for cars or faces as the d' for trials where the top and bottom halves of both stimuli in a pair were the same or different ("consistent" trials), minus the d' for trials where the top halves of stimuli were the same or different but the bottom half of both was different ("inconsistent" trials). For normal cars this "holistic" processing was 0.56 in non-experts, for cars with the top half inverted ("transformed") this was 0.49 in non-experts. The calculations for faces are not available; instead, only a measure of the amount that car processing interferes with face processing is given. This data seems to suggest that transformed cars interfere less with face processing than normal cars (but the results are very difficult to interpret).

Although this task may appear similar to the composite paradigm, there are some very important differences. First, there is no unaligned condition equivalent to that used by Young et al (1987). Although inverting one half instead of off-setting it means that there is less difference between aligned and "transformed" conditions in the visual extent of the stimulus, it also means that instead of two (upright or inverted) half cars (or faces) being presented, half an upright car and half an inverted car are presented. Thus, whereas in Young et al. (1987) the aligned and unaligned conditions are equivalent in all respects except for the relative location of the halves, in Gauthier et al. (2003) they are not. If two things are presented in opposite orientations they are less likely to compete for attention than if both are presented in the same orientation. Thus, the condition in which the top half of the car is inverted (the transformed condition) might be easier than the condition in which both halves are upright without any true holistic processing (i.e., perceptual integration) occurring. Gauthier et al. also did not include a condition with faces transformed in the same manner, making it impossible to say empirically whether the effect for cars resembles an effect for faces.

5.6.4 What is the origin of holistic processing for faces?

Taking the present results and the previous literature together, I have argued that the expertise hypothesis is not supported by the empirical evidence. I therefore suggest that the domain-specificity hypothesis should be accepted as the only currently viable alternative. However, if holistic processing of faces is not explained by the large amount of expertise that we have with faces, an alternative explanation is needed as to the origin of holistic processing only for upright faces; in that domain-specificity is not an explanation in itself.

Infants prefer face-like stimuli to those which are less face-like (M. H. Johnson, Dziurawiec, Ellis, & Morton, 1991; Mondloch et al., 1999; Morton & Johnson, 1991). They are also able to tell their mother's face from a strangers, although in newborns this may be based on hairline (Pascalis, de Schonen, Morton, Deruelle, & Fabregrenet, 1995). By the time infants are 7-months old they integrate information from the internal and external aspects of the face (Cohen & Cashon, 2001). Further, there seems to be a critical period somewhere in the first three to six months for developing holistic processing of faces (Le Grand et al., 2001, 2003, 2004). Thus, it seems likely that either there is some innate basis to face recognition and/or that a period of exposure to faces in early infancy is essential to the normal development of holistic processing.

5.6.5 Objects of expertise and a critical period for developing holistic processing.

I have argued that current research strongly suggests that holistic/face-like processing cannot occur for non-face objects after many years of experience. Importantly, while the expertise gained with dogs by subjects in the present study often began in childhood (as early as four years in one case), it was never gained in early infancy. Does this mean that holistic processing (of the kind that occurs for faces) can never occur for objects-of-expertise?

As noted above, there seems to be a critical period for developing holistic processing of faces. It is possible that this critical period is in fact a critical period for developing holistic processing in general, but that faces are the only things which babies get to see very much. Very few people encourage babies younger than six months to spend extensive amounts of time looking at lots of individual dogs or other non-face objects which share a first-order-configuration. Whether it is possible to develop holistic processing of non-face objects if these were seen extensively in the critical period is not something that can ethically be tested in humans. One possibility would be to train monkeys with individual exemplars of an object class from birth. However, this would first require showing that monkeys use holistic processing for faces (e.g., by demonstrating that they show a composite effect), which has not yet been tested.

Overall, evidence does not support the expertise hypothesis in the current form proposed by Diamond and Carey (1986; also see Gauthier & Tarr, 1997, 2002, etc.) namely that experience gained in late childhood/adulthood is sufficient to produce holistic processing for objects-of-expertise. Extensive further testing is needed before it can be established whether expertise for non-face objects could develop if seen in early infancy during the same critical period as for developing face processing. Regardless of the outcome of such studies, in any practical sense, in normal human infants holistic processing is likely to be restricted to faces.

CHAPTER 6: ARE SOME DIMENSIONS IN FACE-SPACE MORE ADAPTABLE THAN OTHERS?

6.1 Overview

In the present chapter, I turn to a rather different set of questions about the effects of experience than those addressed in previous chapters. Here, I am concerned with the effects of quite short term experience (2 mins of adaptation), and the way in which upright faces are represented as individuals in face-space. I am also interested in whether the effects of short-term experience interact with previous longer-term experience; in particular whether short-term experience interacts with the way in which longer-term experience has tuned the dimensions of face-space to suit the range of faces seen.

Previous researchers have shown that adaptation to distorted faces can be understood in terms of changing the norm of face-space. Here, for the first time, I ask whether different dimensions of face-space might be differently adaptable. Two distortion types associated with different variability in the natural range of inputs to face-space (i.e., different required coding ranges) are contrasted: symmetrically changing the height of both eyes together (large variability/range required) and asymmetrically changing the height difference between eyes (small variability/range required). Results showed that the symmetric distortion produced a greater aftereffect than the asymmetric distortion when distortion level of the adaptor was moderate or extreme with respect to required coding range. Aftereffects were similar in magnitude when the adaptor was within the required coding range for symmetric deviations. Additional results argue that, with the relational (feature spacing) distortions, the aftereffects for upright faces arose largely from the face system: adaptation did not transfer across orientations (upright to inverted, or vice versa). I consider the results in terms of a two-pool broadband neural model for coding dimensions in face-space.

6.2 Introduction.

Adaptation aftereffects occur throughout the visual system, from early vision (e.g., colour processing) to high-level vision (e.g., face recognition). Aftereffects often take the form of a shift in perception in the opposite direction to the adaptor with respect to some perceptual centre point or "norm". For example, in colour vision adaptation to red makes white (the original perceptual centre point on a red-green dimension) appear green. For faces, norm-based aftereffects have been shown for sex, ethnicity and expression (Hsu & Young, 2004; Webster, Kaping, Mizokami, & Duhamel, 2004).

Of direct relevance to manipulations used in the current study, adaptation may also affect face shape. Webster and MacLin (1999; see also MacLin & Webster, 2001) showed that looking at a face contracted around a vertical midline (and/or a horizontal midline) for 2 mins made a normal face look expanded around the midline, and vice versa. Adaptation to an undistorted (normal) face had no effect on perception. Rhodes, Jeffery, Watson, Clifford, and Nakayama (2003) found that adapting to a radially expanded face made an unaltered face appear contracted. In all cases these aftereffects generalised over changes in size and/or position between adaptor and test face (see also Zhao & Chubb, 2001). Rhodes et al. (2003) further demonstrated that adaptation could generalise over task (normality vs. attractiveness judgments). Across studies, methods used to assess aftereffects have included adjusting the face until it looked normal, rating faces for normality/attractiveness, and determining the minimum detectable distortion level in a normal-distorted decision.

In low-level and mid-level vision, adaptation has well established links to neural processing mechanisms (e.g. Clifford, Wenderoth, & Spehar, 2000; Kourtzi & Kanwisher, 2001; Smith & Edgar, 1994). It has been suggested (e.g., Leopold, O'Toole, Vetter, & Blanz, 2001; Rhodes et al., 2003) that adaptation may similarly provide insight into neural mechanisms of face processing. In the present chapter I am concerned with the implications of adaptation studies for understanding the representation of individual faces in "face-space".

6.2.1 Adaptation and face-space.

Face-space is a common heuristic used to explain the coding of face identity (e.g., Valentine, 1991; Valentine & Bruce, 1986). In this approach, each individual face is coded in terms of its position on a series of underlying dimensions. The average face is located at the centre of face-space, and atypical individuals lie towards the periphery. The dimensions of face-space are usually presumed to be built up from experience rather than innately specified.

It is currently unclear which aspects of faces are coded as face-space dimensions. In computational models based on Principal Components Analysis (e.g., Atick, Griffin, & Redlich, 1996; O'Toole, Abdi, Deffenbacher, & Valentin, 1995), the dimensions are the “eigenfaces” which explain variance in an input set of pixel-intensity coded images standardised for face size and orientation. These eigenfaces code information from across the full area of the face, and early eigenfaces (those explaining the most variance between individuals) tend to represent very general variables such as face sex, face fullness, and so on. From behavioural findings (described later), however, one might also suspect that aspects such as spatial relations between face features (e.g., how high the eyes are relative to nose or mouth position) and/or local feature shape information (e.g., the size of the nose) could form useful dimensions of subjects' face-space.

The idea of adaptation within face-space was raised by Leopold, et al (2001; see also Rhodes, Jeffery & Leopold, 2004, described in Rhodes, Robbins et al., in press). Starting from each of four individual faces, Leopold et al. created four corresponding "anti-faces". To do so, a large number of faces were first averaged. The original individual (e.g., Bill) was then morphed on a trajectory towards this average face, with the morphing continued past the average to create the anti-face (i.e., anti-Bill). Theoretically, anti-Bill then lies at an equal distance from the centre of face-space as Bill, but in the opposite direction on all dimensions. For example, if Bill had eyebrows which were 20% more bushy than the average and lips 15% narrower than the average, anti-Bill would have eyebrows 20% less bushy than the average and lips 15% wider. Results showed that, after adapting to these anti-faces for 5 secs, subjects perceived the average face as somewhat like the original individual (e.g., adapting to anti-Bill made the average face look like Bill). This was interpreted in terms of shifting the norm of face-space along the trajectory between the anti-face and the original.

In the present study, the first aim was to develop upon the theoretical link between adaptation and face-space. The Leopold et al. (2001) manipulation varies multiple aspects of facial appearance simultaneously (face width, lip thickness, distance between the eyes, etc). I was interested in whether it would be possible to observe adaptation for distortions that change a single aspect of the face at a time (e.g., shifting the eyes while leaving the rest of the face unaltered), and which might correspond to relatively simple dimensions in face-space. Further, I wished to ascertain whether some dimensions in face-space might be more sensitive to adaptation than others – that is, whether it is easier to temporarily shift the norm of face-space in some directions than in others – and whether this would relate to the range of faces which need to be coded on given dimensions.

6.2.2 Contrasting two dimensions with different variability in face-space.

Both manipulations used here involved changing eye height within the face. Adaptation was contrasted for two different forms of eye-height distortion, differing in the range of stimulus values that face-space might be expected to code.

The two types of changes contrasted are illustrated in Figure 6.1. In the symmetric eye-height distortion, faces were altered by moving the two eyes up together ("positive" direction), or the two eyes down together ("negative" direction). An adaptation aftereffect for this distortion would manifest as, for example, the original undistorted face appearing to have its eyes too low after adapting to an eyes-up face. Readers can test this phenomenon for themselves in the figure. In the asymmetric eye-height distortion, the alteration involved moving either the right eye up and the left eye down ("positive" direction), or vice versa ("negative" direction). The amount of distortion in each face was defined by the amount of change per eye; thus, a "50 pixel-per-eye" symmetric distortion (e.g., 50 pixel right eye shift up, and 50 pixel left eye shift up) was matched for total metric deviation to the "50 pixel-per-eye" asymmetric distortion (e.g., 50 pixel right eye shift up, and 50 pixel left eye shift down)[1].

---

[1] 50 pixels in these stimuli corresponds to approximately 19% of the hairline to chin distance.

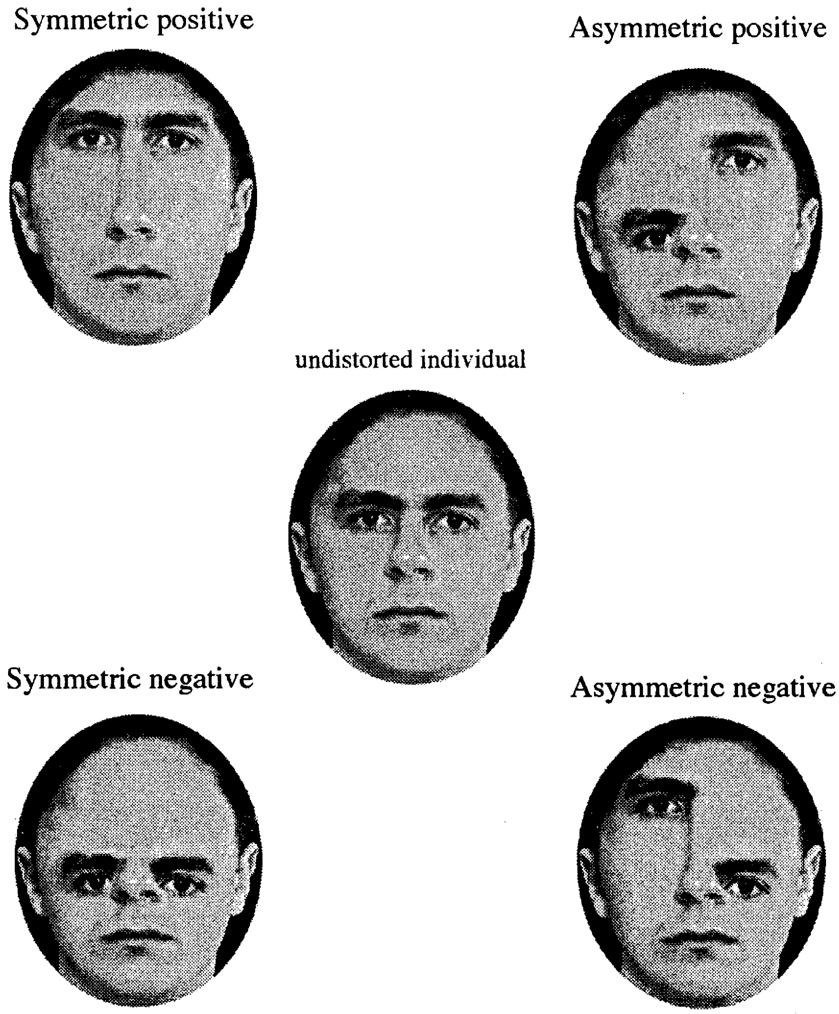


Figure 6.1. Examples of the most extreme distortions (±50 pixels-per-eye) on Bill, as well as an undistorted face (centre). The effect of adaptation can be observed by scanning one of the distorted faces for 30 s and then looking at the undistorted face.

The symmetric and asymmetric eye-height distortions were chosen because they correspond to aspects of faces that have different amounts of variability in real world images, and thus might be expected to require a different coding range of eye-height values in face-space. Physical measurements of faces show that the mean height of both eyes together in the face (i.e., symmetric changes) varies quite considerably across individuals. For Caucasians, the distance of eyes to hairline in young adult males has a standard deviation of 7.5 mm (around a mean of 67.1 mm), and the distance from eyes to chin has a standard deviation of 5.7 mm (around a mean of 124.7 mm; Farkas, Hreczko, & Katic, 1994). Further, up-down head rotation (e.g., nodding) also substantially changes the apparent position of eyes with respect to other features. In contrast, the asymmetry between eyes varies to a much smaller degree. Although 69%

of the adult males measured by Hreczko and Farkas (1994) had some asymmetry between the height of their left and right eyes, the standard deviation was only 1.1 mm (around a mean difference of 2.8 mm). There is also no three-dimensional head movement that can increase the physical asymmetry between the eyes (e.g., tilting the head sideways leaves the distance from each eye to the hairline unchanged). Presuming that face-space is tuned to efficiently code the range of inputs to which it has been exposed, I would then expect that the dimension associated with symmetric eye-height differences would code a large range of eye-height values, while the dimension associated with asymmetric eye-height differences would code a smaller range of eye-height values [2].

A previous study (McKone, Aitkin, & Edwards, submitted) has tested subjects' perception of symmetric and asymmetric eye-height distortions in the unadapted state. For the symmetric distortion, subjects judged a relatively broad range of eye heights around the original as all "normal" or "just like Bill". For the asymmetric distortion, subjects accepted a much smaller range of eye heights as "normal" or "just like Bill". Subjects gave a normality or identity rating across 21 levels of distortion ranging from -50 to 0 to +50 pixels-per-eye. As face deviation increased away from 0 (in either direction), subjects initially perceived no change in normality/identity from the undistorted face. The critical finding was that the width of this "below threshold" region was substantially greater for symmetric eye-height distortions (4.89 pixels away from zero in each direction) than for asymmetric eye-height distortions (1.79 pixels in each direction). Outside this range, ratings were related to physical deviation in a logarithmic fashion. That is, for a unit increase in deviation, changes in perception were large immediately past the threshold, but were smaller towards extreme deviation levels (this theshold-plus-log curve shape is known as Fechner's law, an extension of Weber's law [3]). The slope of the logarithmic section was shallower for symmetric than asymmetric distortions, again indicating weaker sensitivity to deviations in the former case.

McKone et al. (submitted) argued that these results could be interpreted in terms of different variability, and thus different coding ranges, associated with different

[2] By the term "dimensions" I do not necessarily mean to imply that these are "core dimensions" or "cardinal axes" (see General Discussion); instead, readers should think of them merely as a line passing through the centre of face-space.

[3] A standard example of Fechner's law (1966, 1860) from early vision is the relationship between physical and perceived light intensity. Here, a very small amount of light added to complete darkness is not perceived at all (the below-threshold region). Once above threshold, perceived intensity increases logarithmically with physical intensity (e.g., lighting a single candle has more perceptual effect in a semi-darkened room than in bright sunlight).

dimensions in face-space. This is illustrated in Figure 6.2. Presuming that subjects' face-space is sensitive to the range of previously experienced faces, exemplar faces should be more variably placed on a height-of-both-eyes-together line drawn through face-space than on a (presumably orthogonal) line of height-difference-between-eyes. To explain the existence of a subthreshold region, McKone et al. argued that normality/identity judgements are made taking this variability into account; for example, for a face to be perceived as "normal", it does not have to be exactly average, but instead must fall within the normal range.

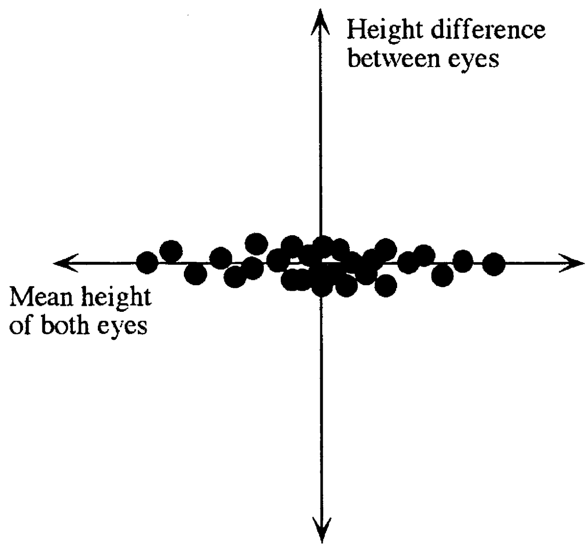


Figure 6.2. Distributions of exemplars (black dots) in face-space along dimensions corresponding to symmetric distortion (mean height of both eyes) and asymmetric distortion (height difference between eyes). Note the greater variability in exemplar placement for symmetric (after McKone et al., submitted).

6.2.3 Does it matter where the adaptor is placed?

If a larger range of values is coded in face-space for symmetric than asymmetric eye height, a larger range of values might also be adaptable for symmetric than asymmetric distortions. Previous studies have tested only one adaptor position. In particular the stimulus used as the adaptor has been the most extreme deviation level within the stimulus set. The second aim of the present study was to test several adaptor positions, corresponding to different distances from the unadapted norm.

A standard model of norm-based adaptation (see General Discussion) would predict that aftereffects should be greater at more extreme values within the coded range, but that outside that range there would be no adaptation. If the range of values coded is greater for symmetric than asymmetric eye height, there might thus be adaptor positions for which the aftereffect is larger for symmetric than asymmetric distortions. Thus, predictions can be made both about the size of the aftereffect at a particular adaptor position, and about the range of adaptor positions which will produce an aftereffect.

6.2.4 How much of adaptation for faces comes from the face-specific system?

Adaptation to distorted faces can arise from multiple stages of the visual processing stream. These include early-, and mid-level vision in addition to both face and object areas within high-level vision. The third aim of the present study was to demonstrate that the adaptation observed had an origin within the face recognition system, at least when the faces were upright.

A contribution to the aftereffect from early retinotopic areas is usually ruled out by changing size and/or retinal location between adaptor and test. Watson and Clifford (2003) also showed that adaptation was not retinotopic by showing that adaptation to faces expanded or contracted along the vertical midline followed the orientation of the face stimulus (45° to the right versus 45° to the left) rather than the absolute direction of the distortion.

A contribution from mid-level shape-processing is more difficult to exclude. Adaptation has been demonstrated for several simple shape properties. For aspect ratio (i.e., direction of elongation), Regan and Hamstra (1992) showed that adapting to a vertically stretched rectangle made a square appear horizontally stretched. This effect transferred across size and also shape type – adapting to a vertically stretched rectangle made a circle appear as a horizontally flattened ellipse – indicating that the effects originated from post-retinotopic areas of visual cortex. Adapting to a square (i.e. the perceptual centre point) had almost no effect on perception. For other shape properties, Suzuki and Cavanagh (1998) showed that adapting to a left pointing triangle made a square appear as a trapezoid pointed in the opposite direction, and adapting to an upwardly curving shape resulted in a diamond being perceived as curving downward. Suzuki (2001, 2003) also showed that adaptation to a concave hour-glass shape caused a set of four squares (positioned as the corners of a larger square) to be perceived as tilted into a convex shape. These effects were again robust across changes in size. For faces, therefore, some component of shape adaptation could arise through adaptation of these mid-level mechanisms. For example, if adapting to a face with widely-opened eyes made the normal face appear to have narrow eyes, this could potentially be attributed to adapting an aspect ratio mechanism.

Turning to high-level vision, it is also important to rule out adaptation arising from non-face object recognition area/s. Functional MRI studies show a face-selective "Fusiform Face Area" (FFA) in inferotemporal cortex, which is more responsive to faces than to a wide range of other objects (Grill-Spector, Knouf, & Kanwisher, 2004; Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Allison, 1997), and is associated with processing individual face identity (e.g., Grill-Spector et al., 2004; Rhodes, Byatt, Michie, & Puce, 2004). However, faces stimuli can also produce some response in areas that are primarily responsive to human bodies, isolated face parts, houses, and other non-face objects (Downing, Jiang, Shuman, & Kanwisher, 2001; Tong, Nakayama, Vaughan, & Kanwisher, 1998), and these areas are also potential sources of adaptation.

In the present study, both the symmetric and asymmetric distortions used were relational in nature. That is, both distortions altered the spacing between face features without (so far as possible) making changes to local feature shape. This type of distortion was chosen to increase the chances of targeting the high-level face recognition system, at least when the face was upright.

To ensure this aim had been achieved, transfer of adaptation across upright and inverted orientations was tested. An extensive literature argues that the processing of inverted faces is different from that of upright faces (e.g., Tanaka & Farah, 1993; Yin, 1969; Young, Hellawell, & Hay, 1987). Of particular relevance here is that sensitivity to relational information in faces is noticeably weaker for inverted faces than for upright faces. If a spacing change is quite small, it can fail to affect subjects' perception of inverted faces at all, despite producing a clear effect on perception for upright faces (e.g., Gilchrist & McKone, 2003; Leder & Bruce, 1998). If a spacing change is larger (e.g., shifting the mouth halfway towards the chin), then it will affect perception of inverted faces, but still more weakly so than in upright faces (e.g., Bartlett & Searcy, 1993; Le Grand, Mondloch, Maurer, & Brent, 2001; McKone et al., submitted). In contrast to the large effects of orientation on spacing changes, sensitivity to local feature changes (e.g., thickening the eyebrows, or changing the eye shape) is commonly almost unaffected by inversion (Bartlett & Searcy, 1993; Gilchrist & McKone, 2003; Le Grand, Mondloch, Maurer, & Brent, 2001; Leder & Bruce, 1998; McKone et al., submitted).

A strong way to show that an adaptation effect for upright faces is truly face-specific is to demonstrate that different populations of neurons support the adaptation effect for upright faces from those supporting adaptation for inverted faces. This could be shown in two ways: demonstrating a lack of transfer of adaptation between upright and inverted faces, or demonstrating that upright and inverted faces can simultaneously be adapted to different competing distortions.

Rhodes, Jeffery, Watson, Jaquet, Winkler and Clifford (2004) have reported the latter result. They simultaneously adapted upright and inverted faces to <u>different</u> global distortions (e.g., upright to expanded and inverted to contracted). They also adapted upright and inverted faces to different sexes (e.g., upright to female and inverted to male). The critical result was that concurrent but opposite aftereffects were elicited in the two orientations. This provides good evidence that upright and inverted faces can be processed by different populations of cells.[4]

Transfer of adaptation between upright and inverted faces has given less clear results. Two studies (Watson & Clifford, 2003; Webster & MacLin, 1999) have crossed the orientation of the adaptor (Upright vs. Inverted) with the orientation of the test stimuli (Upright vs. Inverted). Both studies found transfer of adaptation across orientations. From upright adaptors to inverted test faces (U-I), adaptation approached the strength of upright adaptors to upright test faces (U-U). There was a weaker effect for inverted adaptors to upright test faces (I-U), but this was still above zero. Thus, in contrast to the concurrent opposite aftereffects (Rhodes, Jeffery et al., 2004), these previous results do <u>not</u> clearly dissociate neural populations for upright and inverted faces, and thus allow that much of the adaptation observed might have come from mid-level vision and/or from object processing mechanisms rather than face-specific mechanisms.

The type of distortions used in previous studies – midline expansion or contraction (Webster & MacLin, 1999; Watson & Clifford, 2003), radial expansion or contraction (Rhodes et al., 2003), and anti-faces (Leopold et al., 2001) – cannot be claimed to specifically target the face system for upright faces. All of these distortions change relational properties of the face (e.g., distance between the eyes) but also produce substantial changes in local feature shape (e.g., making the eyes rounder or

---

[4] This might appear to conflict with evidence that the FFA shows very little difference in response to upright and inverted faces (Kanwisher, Tong, & Nakayama, 1998). However, it is possible that the FFA includes distinct populations of neurons which code upright and inverted faces, and this information is lost due to insufficient spatial resolution of the blood oxygen level dependent (BOLD) response.

narrower, or turning down the corners of the mouth). Given the evidence of particularly strong sensitivity to relational changes in upright faces, in the present study, I introduce relational rather than global (relational-plus-local) distortion types, with the aim of demonstrating no transfer of adaptation across orientations. Such a result would argue that the adaptation observed for upright faces was genuinely face-specific.

Note that I did not expect to find zero adaptation for inverted faces with inverted adaptors (I-I condition). Relational changes can be perceived in inverted faces if they are large enough, and adaptation for inverted faces (I-I) has been strong in previous adaptation studies (Leopold et al., 2001; Watson & Clifford, 2003; Webster & MacLin, 1999). This adaptation does not necessarily come from the face system, and may arise from some other source. To show a face-system origin of upright face adaptation, the relevant result to demonstrate would be the lack of transfer between orientations.

6.2.5 Structure of the present experiments and general methodological approach.

In Experiment 10, I contrasted different dimensions of face-space, using only upright faces. The primary question was whether the distortions along the dimension of symmetric eye-height (associated with high variability/lage coding range in face-space) would produce more adaptation than distortions along the dimension of asymmetric eye-height (associated with low variability/small coding range in face-space). In this experiment, the adaptor was positioned to be moderately extreme with respect to the required coding range for both distortion types. In Experiment 11, again testing only upright faces, the position of the adaptor relative to the norm was varied to assess the effects of symmetric versus asymmetric distortions at different positions within their required coding ranges. In Experiment 12, I examined transfer of adaptation across orientations (i.e., in the I-U and U-I conditions) with the aim of showing that aftereffects for upright faces arose from adaptation of face-specific mechanisms.

In each experiment, the task was either to rate each stimulus for "how much like" a particular person it appeared, or to adjust the stimulus to appear "most like" a particular person. The judgment involved reference to identity, rather than merely normality, to encourage processing of the stimuli as individual faces (i.e., a within-class discrimination task rather than a between-class discrimination task). To familiarise subjects with the target identities, each experiment began with a training phase in which subjects learned to recognise four individuals (Bill, Sam, John and Fred; see Figure 6.3) in their undistorted form.

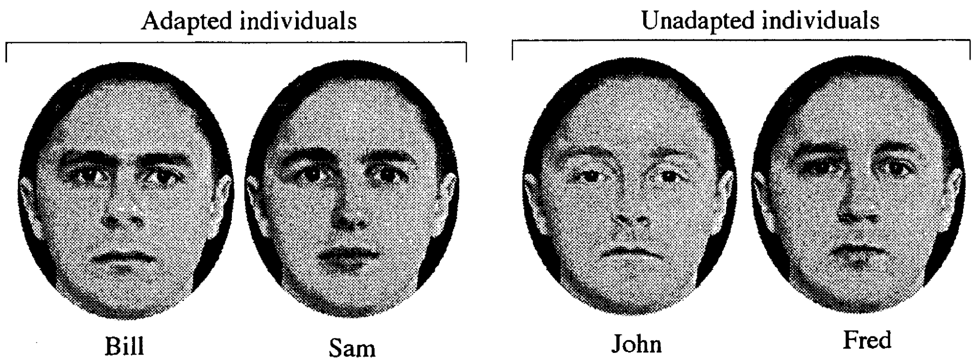


Figure 6.3. The four individuals used in the experiments, shown undistorted.

To assess adaptation aftereffects, a three-phase procedure was used for each distortion type (symmetric or asymmetric). The pre-adaptation phase presented multiple levels of the relevant distortion type for rating, or asked the subject to adjust the face, to allow determination of the baseline image perceived as "most like Bill/Sam/John/Fred". The adaptation phase then presented a particular distortion level of the relevant type for 2 mins. The post-adaptation phase was identical to the pre-adaptation phase, with the addition of top-up adaptation stimuli presented for 5 s between each of the test stimuli.

To avoid a contribution of low-level retinotopic aftereffects, the adapting faces were presented at a smaller size than the test faces. Similarly, subjects were instructed to make eye movements around the adaptor during the adaptation phase, and top-up adaptor faces were presented at several locations to require eye movements during the post-adaptation test.

Another important point was that, of the four individuals used as test faces (Bill, Sam, John, Fred), only two were used as adaptors (Bill and Sam). This allowed, for the first time, a full examination of transfer across identity by contrasting the amount of adaptation when the adaptor and test identities were the same (i.e., for Bill and Sam) with the amount of adaptation when they were different (i.e., for John and Fred). Rhodes et al. (2003) showed nonzero transfer across identity, but tested only a different-identity condition. In the present experiments, strong transfer across identity would argue that adaptation aftereffects can be explained as a general shift in the norm of the relevant face-space dimension. This idea is illustrated in Figure 6.4. Alternatively, a finding of weak or no transfer across identity would suggest that adaptation shifts the position of an individual face within face-space, rather than shifting the norm of the entire dimension.

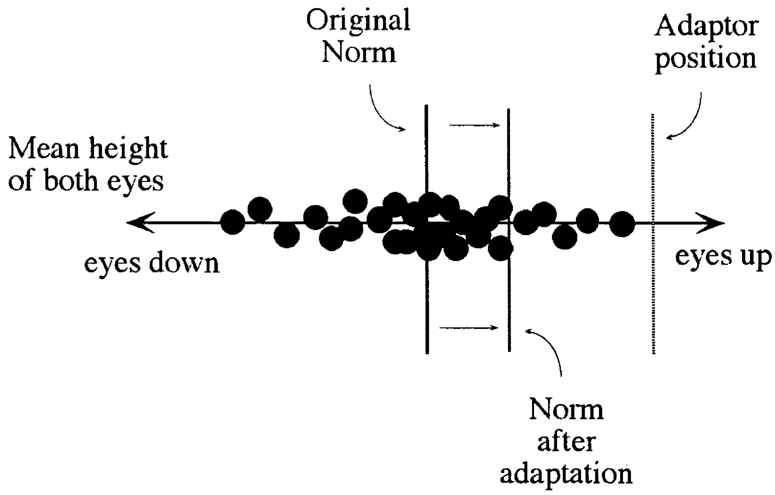


Figure 6.4. Adaptation conceptualised as a simple shift in the position of the norm of a dimension in face-space, illustrated using the height of both eyes together dimension.

6.3 Experiment 10: Adapting Different Relational Dimensions in Face-Space

In this experiment the two relational distortions (symmetric and asymmetric) were used to test the idea that different dimensions in face-space might be differently adaptable. Analysis was based on single-subject data. Each subject rated pictures of Bill, Sam, John and Fred at 27 levels of distortion (ranging from -50 to +50 pixels-per-eye) for how much the stimulus looked like the original person. This was done for each distortion type (for a total of 2.5 hours per subject). Only upright faces were tested and only one adaptor position was used.

The adaptor position was set at an intermediate level of distortion (±20 pixels-per-eye). All previous studies have used the most extreme test stimulus as the adaptor. One advantage of using an intermediate value is that it allows perceptual changes to be assessed on both sides of the adaptor (i.e., for more extreme deviation levels as well as less extreme deviations). In fact, three regions are of interest: faces more extreme than the adaptor on the adaptor side of the original norm (for a +20 adaptor, this is +21 to +50 pixels); faces falling between the adaptor and the original norm (0 to +19 pixels); and faces falling on the opposite side of the original norm to the adaptor (-50 to -1 pixels). If an aftereffect arises from a general shift in the norm of a dimension in face-space (Figure 6.4), the prediction is that faces on the adaptor side of the original norm, regardless of which side of the adaptor they lie on, should be perceived as more like "Bill" than previously, while faces on the non-adaptor side of the original norm should be perceived as less like "Bill" than previously.

6.3.1 Experiment 10 - Method

6.3.1.1 Subjects.

There were five subjects in total (age range 20-42; two male), four naive as to the purposes of the experiment (S1-S4), and the experimenter (RR). All subjects were Caucasian, and thus ethnically matched to the face stimuli. All reported normal or corrected-to-normal vision. Naïve subjects were paid $10 per session. For the four naïve subjects, testing was interleaved with some conditions of Experiment 12.

6.3.1.2 Design.

For a given distortion type (e.g., symmetric), ratings of how much a deviated face looked liked the original individual (1 = "looks exactly like" Bill/Sam/John/Fred, 9 = "looks nothing like") were determined for 27 deviation levels (0, ±1, ±2, ±3, ±6, ±8, ±10, ±12, ±15, ±20, ±25, ±30, ±40, and ±50 pixels-per-eye), for all four individuals (Bill, Sam, John and Fred). After adaptation to either +20 or -20 adaptors (Bill and Sam) of the same distortion type, ratings were taken again for all four individuals. "Adapted" individuals were thus Bill and Sam, and "unadapted" individuals were John and Fred.

Across different sessions, this procedure was repeated for symmetric distortions with a positive adaptor (eyes up), symmetric distortions with a negative adaptor (eyes down), asymmetric distortions with a positive adaptor (right up, left down), and asymmetric distortions with a negative adaptor (right down, left up). The ±20 adaptor stimuli are illustrated in Figure 6.5. Each session lasted 30 – 45 mins. Subjects usually completed only one session per day, with a minimum time between sessions of 30 mins. Order of conditions was counterbalanced across subjects.

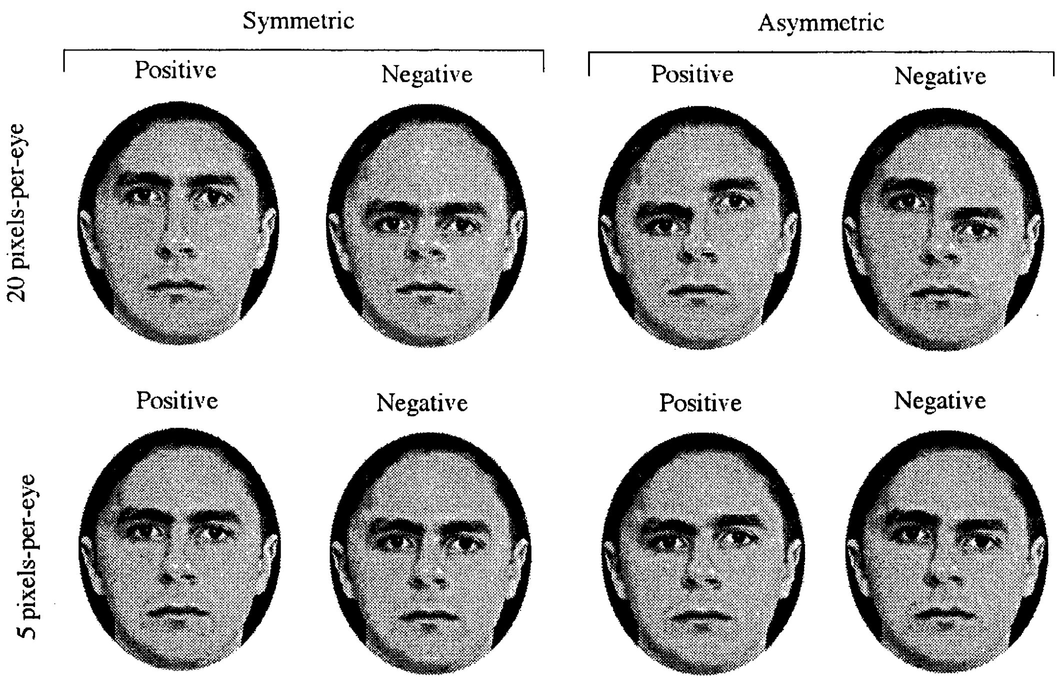


Figure 6.5. Adaptor deviations used in Experiments 10 & 12 (±20), and Experiment 12 (±5 and ±20; the ±50 is shown in Figure 6.1), illustrated using Bill.

6.3.1.3 Stimuli.

All stimulus manipulations were carried out with Adobe Photoshop 5.5. Stimuli were derived from four undistorted 0 deviation faces taken from McKone et al. (submitted; see Figure 6.3), and labelled as Bill, Sam, John and Fred. These had been made by cutting the internal regions from photographs of four real people (obtained from the Stirling PICS database), and pasting them onto a common background head derived from a fifth person. The natural distances between eyes, nose and mouth were retained, and the overall set of internal features was placed at an average height within the new head. A common background head with stiff upright hair was used to ensure that eye/s could be moved relative to a clearly-seen hairline. The resulting faces were all natural in appearance.

From each of the four undistorted individuals, 27 deviation levels were made for each distortion type (symmetric and asymmetric). The total number of pixels from hairline to chin in the images was 263; thus the most extreme deviation of 50 pixels-per-eye corresponded to 19.01% of the hairline-chin distance, the adaptor position of 20 pixels-per-eye to 7.60%, and 1 pixel-per-eye was equal to 0.38% of the hairline-chin distance. At the viewing distance of 40 cm, test stimuli (i.e., those shown in the pre-adaptation and post-adaptation phases) subtended 10° from the chin to hairline by 7.9°

across the widest point of the cheeks (and thus a 10 pixel-per-eye deviation, for example, corresponded to 0.4° of movement if fixating the original eye position). Adaptor stimuli were shown smaller than the test stimuli, at 7.9° by 5.7°.

6.3.1.4 Procedure.

Stimuli were presented on an iMac computer with a 36 cm screen set to a resolution of 1024 x 768 pixels, using PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). A chin-rest was used.

Training phase. An identity training phase began with each 0 deviation face (Bill, Sam, John and Fred) shown once upright and once inverted[5] with its name presented beneath. This was followed by 32 trials of training-with-feedback (four trials of each 0 deviation face in each orientation). Here, stimuli were presented without names, subjects pressed the key matching the face's initial (B, S, J, or F), and a feedback beep was given if the response was incorrect.

At the beginning of their first session, each subject completed four cycles through this training phase (i.e. 40 exposures to each individual), taking approximately 15-20 mins. Identification accuracy was measured during the training-with-feedback task, and all subjects passed a criterion of 80% correct in the final cycle. At the beginning of each additional session, a single cycle through the training phase was given to ensure subjects had reinstated an active representation of each face in its undistorted form.

Pre-adaptation test phase. In the pre-adaptation phase of each session, each of the four individuals (Bill, Sam, John and Fred) was presented once at each of the 27 deviation levels, for a total of 108 trials. Face stimuli were presented at the centre of the screen with no fixation point (this avoided providing easy cues as to absolute location of the eyes on the screen, rather than within the face). On each trial, the stimulus face was shown for 1000 ms, followed by a 200 ms blank screen, and then a prompt to enter the identity rating. The name of the person was always given in the prompt. For example, following a picture of Bill, the question read "How much does this look like Bill? 1= 'exactly', 9= 'not at all'". Trials were presented in random order for each subject.

[5] The training phase included inverted faces as well as upright faces because some conditions of Experiment 12 were interleaved with those of Experiment 10.

To encourage appropriate use of the full scale, subjects were shown examples from the relevant distortion type (symmetric or asymmetric) before the rating task began. Six examples covering both mild and extreme deviation levels in both directions were shown, using a nontested individual (Ralph).

Adaptation phase. In the adaptation phase of each session, the relevant adaptor deviation (e.g., 20 pixels eyes up, symmetric distortion) was shown for 2 mins. The adaptor individual was Bill for 1 min then Sam for 1 min, or vice versa, with the individual presented first chosen randomly. Adaptors were presented in the centre of the screen. Subjects were instructed not to fixate the adaptor, but instead to scan the whole face for the full exposure time.

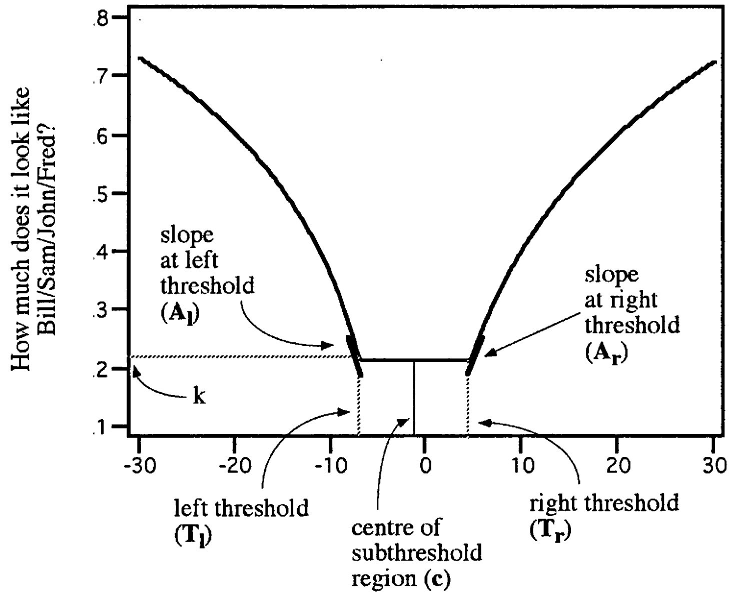
Post-adaptation test phase. The post-adaptation phase in each session was a repeat of the pre-adaptation phase, with the addition of top-up adaptation. Before each rating trial began, one of the adaptor individuals (Bill or Sam) was shown for 5000 ms, with alternation of adaptor identity on successive trials. The top-up adaptor was shown in one of four locations around the test face, to force subjects to make eye movements to view the adaptor. The locations were centred 5.7°(vertical) and 5.7° (horizontal) away from the test face, to the top-left, top-right, bottom-right or bottom-left corner, with locations chosen in clockwise order on successive trials.

6.3.2 Experiment 10 – Results

6.3.2.1 Curve fits.

To allow quantitative summary of the data in any given condition, curve fitting was carried out. The primary analysis was based on fits using Fechner's law (threshold-plus-log curves). As explained in Figure 6.6, Fechner law fits allow extraction of five parameters of interest. (In reading all graphs, note again the direction of the identity rating scale: 1= most like "Bill"; 9 = least like "Bill".) The left threshold value (T_l) indicates the number of pixels-per-eye by which the stimulus can be deviated in the negative direction before ratings of identity start to suddenly increase. The right threshold value (T_r) is the corresponding value for positive deviations. The centre of the full subthreshold region (c) provides the best estimate of the stimulus perceived as "most like" the individual. The final two parameters are the left slope-at-threshold (A_l)

and the right slope-at-threshold (A_r); in each case, higher slopes indicate greater sensitivity to changes, this time in the above-threshold region. All fits were done using CurveFit (Kevin Raner Software)[6].



$$y = \begin{cases} A_l \ln(z-x) + k, & \text{for } x < T_l \\ k, & \text{for } T_l > x < T_r \\ A_r \ln(x-z) + k, & \text{for } x < T_r \end{cases}$$

where z is the common vertical asymptote for the two log sections.

Figure 6.6. The Fechner law function used to fit the data.

In the pre-adaptation conditions, Fechner law curves provided good fits to every subject's data. Figure 6.7a shows an example of a fit for a single subject in a symmetric distortion condition. For data in the asymmetric condition, the Fechner curves sometimes failed to capture the full degree of “pointedness” close to the zero deviation, as illustrated for one subject in Figure 6.7b. For this reason I also tried fitting Pearson IV curves. These are 4-parameter modified Gaussians that allow more pointedness (kurtosis) than the standard 3-parameter Gaussian. Although the overall quality of the fits was still good for Pearson IV, average R^2 did not increase compared to the Fechner

[6] This program allows the threshold points to be determined from the data, unlike the majority of programs which require them to be pre-specified. It does not allow different line patterns for fits.

fits, and the ability to capture the pointedness did not improve (Figure 6.7c). I chose to use Fechner law fits, because they do not force the curve to be symmetrical (thus providing better fits in cases like Figure 6.7a), and because there is a well-established theoretical rationale behind Fechner's law. Further, Fechner's law provided good fits in McKone et al. (submitted) where there were considerably more subjects. Conclusions regarding adaptation were the same from both fitting methods. To demonstrate that the good fits with the 7-parameter Fechner law could not be attributed simply to the number of parameters, Figure 6.7d shows that fits from a 7th-order polynomial were poor.

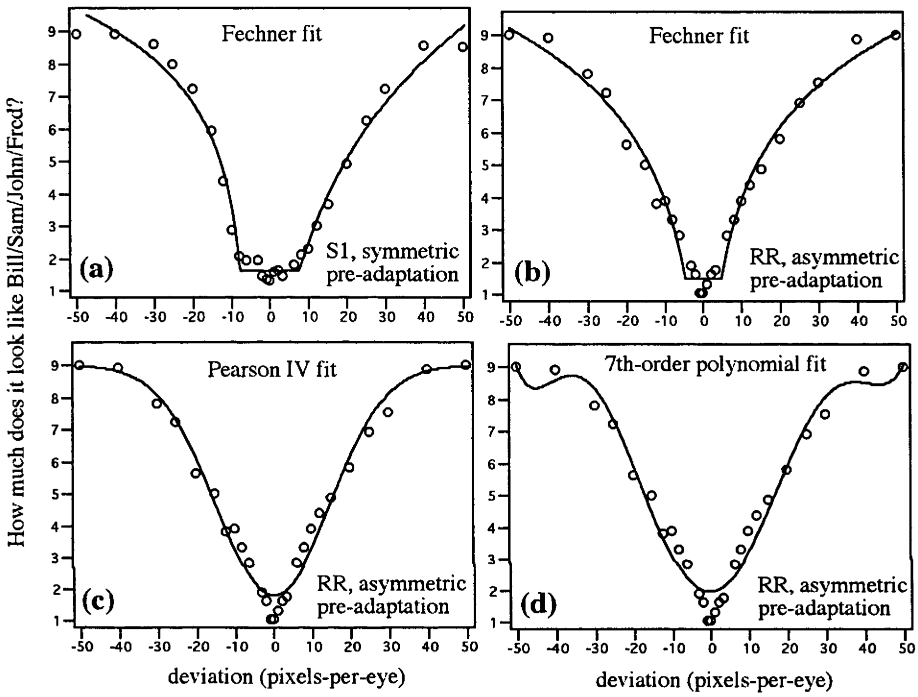


Figure 6.7. Examples of fits, for single subject data in single conditions. Fechner law fits are shown in (a) and (b) for two cases. Panels (c) and (d) show the same data as (b) to illustrate that neither a Pearson IV curve nor a 7th-order polynomial improved the fit.

6.3.2.2 Confirming greater perceived variability for symmetric than asymmetric distortions.

Before analysing adaptation effects, it was important to examine the pre-adaptation data to check that, for the current subjects, the range of faces perceived as normal was larger for the symmetric distortion than for the asymmetric distortion, thus suggesting a greater coding range in the former case (cf. McKone et al., submitted). To give the most reliable results, I averaged over all relevant unadapted scores. This included collapsing over direction of future adaptor (+20, -20), identity of the face (Bill, Sam, John and Fred were all equivalent here since no adaptation had yet taken place), and also adding in extra trials from the pre-adaptation conditions obtained for upright faces in Experiment 12. (Note in advance that all calculations comparing pre- and post-adaptation used data only from Experiment 10).

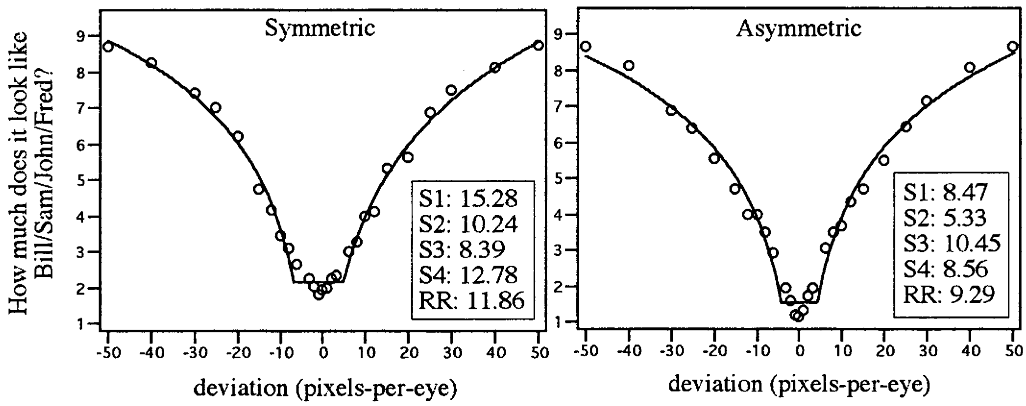


Figure 6.8. Experiment 10: Pre-adaptation identity ratings (averaged over subjects, as well as identity of face and direction of future adaptor) and the Fechner law fit to the average data. The inset box indicates the total width of the sub-threshold region ($|T_l| + |T_r|$, in pixels-per-eye) determined for separate fits, for each individual subject.

Figure 6.8 shows the resulting ratings and fitted Fechner curves for the symmetric distortion (Figure 6.8a) and the asymmetric distortion (Figure 6.8b), for data averaged over the five subjects. In the lower right corner of each plot, the width of the subthreshold region ($|T_l| + |T_r|$) is also listed for each individual subject. The key result is that, as in McKone et al. (submitted), subjects were less sensitive to symmetric deviations of eye height than to asymmetric deviations. The subthreshold region was consistently wider for symmetric than asymmetric distortions; that is, for symmetric distortions, subjects tolerated more physical deviation in the face while still considering it as much like Bill as the original face. These results confirm that the current subjects' face-space includes greater variability associated with the symmetric eye height dimension than the asymmetric eye height dimension, corresponding to the greater variability within the experienced inputs to face-space in the former case.

6.3.2.3 Adaptation effects.

The primary theoretical questions were then about adaptation effects. Figure 6.9 shows data averaged across subjects, for pre- and post- adaptation, broken down by direction of adaptor (+20, -20) and identity status (adapted individuals = Bill and Sam; unadapted individuals = John and Fred). Three important results are apparent. First, adaptation took the form of a simple shift in ratings in the predicted direction. Second, adaptation shifts were larger for symmetric than asymmetric distortions. Third, adaptation generalised completely over adaptor-test changes in identity.

Form of adaptation. The presence of a norm-based adaptation aftereffect is indicated by a shift in the rating curve towards the adaptor direction. To make this prediction clear, consider that, after adapting to an eyes-up face (+20), the unaltered face (0) should appear to have its eyes moved down; thus, the face which is perceived as "most like" the original will be a stimulus with its eyes physically higher than the original position. This corresponds to the centre of the subthreshold region shifting in the positive direction, for a positive adaptor.

Figure 6.9 shows that such adaptation occurred. For example, after adapting to the +20 symmetric face in Figures 6.9 (right panels), stimuli with eyes deviated in the positive direction (eyes up) now appeared more like Bill/Sam/John/Fred (i.e., lower ratings) than they did prior to adaptation, while stimuli with eyes deviated in the negative direction (eyes down) now appeared less like Bill/Sam/John/Fred (i.e., higher ratings). This pattern – that is, ratings decreasing on the adaptor side of the original zero point, while increasing on the other – produces an overall shift in the curve towards the direction of the adaptor. It can be seen from Figure 6.9 that adaptation was equally strong for both directions, producing a rightward shift for positive adaptors (collapsed across identity, symmetric = 6.39 pixels; asymmetric = 1.03), and a similar-sized leftward shift for negative adaptors (symmetric = 5.35 pixels; asymmetric = 1.25).

Also note that faces more extreme than the adaptor (e.g. the +30 stimulus) were affected in the same way as faces lying in the region between the adaptor and the original zero stimulus (0 to +19), in that all faces whose deviations levels were above threshold now appeared more like Bill than before adaptation. This pattern of a simple overall shift is consistent with a re-centring of the norm of the relevant dimension in face-space (Figure 6.4).

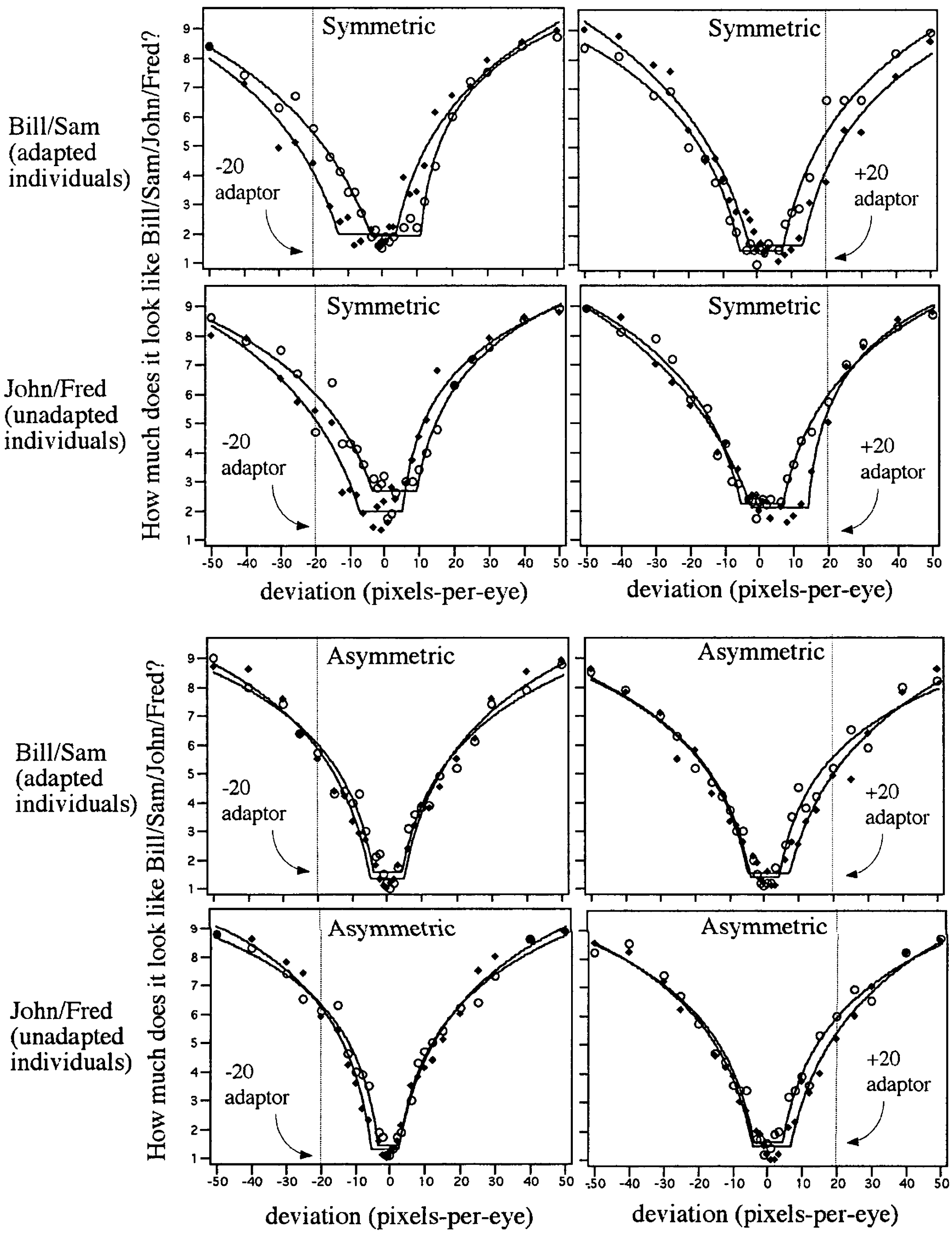


Figure 6.9. Experiment 10: Identity ratings for the pre-adaptation condition (open circles) and post-adapted condition (filled diamonds), plus Fechner law fits, averaged over subjects.

Symmetric versus asymmetric distortions. The major result apparent in Figure 6.9 is that the amount of adaptation was greater for symmetric distortions of eye height (top four panels) than it was for asymmetric distortions (bottom four panels). That is, with the adaptor positioned at ±20 pixels, the amount of shift in the adaptor direction (rightwards shift for positive adaptors, leftwards shift for negative adaptors) was consistently greater for the symmetric condition.

To allow concise presentation of individual subject data, I calculated a simple measure of shift, as the difference between the centre of the subthreshold region before and after adaptation. Shift scores were defined such that a positive shift always corresponded to a change in the direction predicted for adaptation (i.e., for positive adaptors, $shift = c_{post} - c_{pre}$; for negative adaptors, $shift = c_{pre} - c_{post}$).

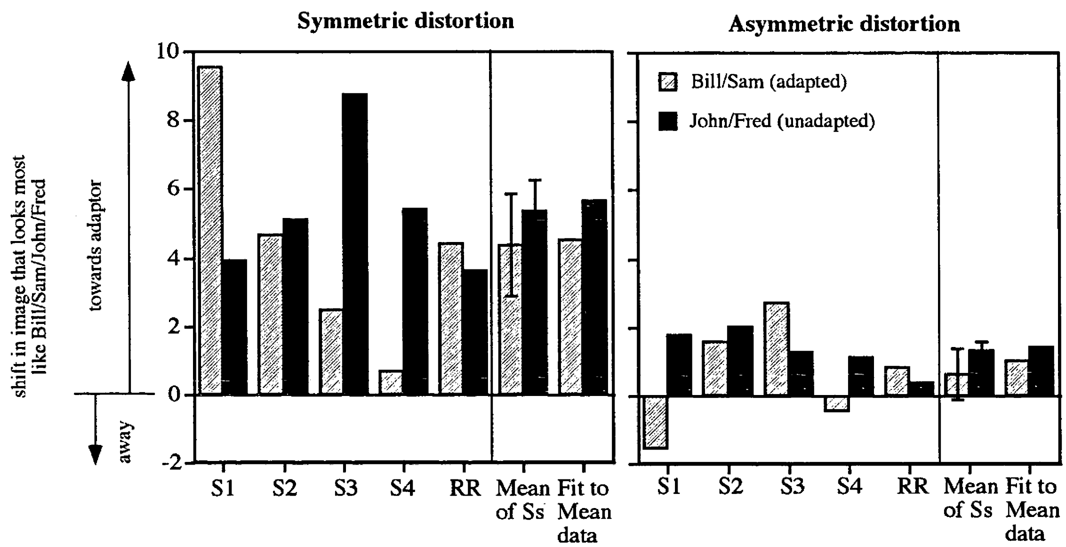


Figure 6.10. Experiment 10: Adaptation results for individual subjects (number of pixels shift in the centre of the sub-threshold region, between pre- and post-adaptation). Results are collapsed over direction of adaptor (positive or negative) as described in the text. Each plot also shows the mean (± 1 SEM) of these individual subjects' scores (denoted "Mean of Ss" on the figure), as well as a score derived by averaging data across subjects first and then fitting (i.e. "Fit to mean data").

Figure 6.10 shows shift scores for the five individual subjects. To increase reliability, individual subject fits were carried out collapsed over direction of adaptor (after first flipping the x-axis for negative adaptor conditions). It is apparent that the greater adaptation effect for symmetric than asymmetric distortions was consistent across subjects. Indeed, the difference between these conditions (collapsed over identity as well as direction) was significant even with only five participants, $t(4)=3.19$, $p<.05$.

Generalisation over identity. Figure 6.9 (full curves averaged over subjects) and Figure 6.10 (shift scores for individual subjects) allow comparison of adaptation for the adapted individuals (Bill and Sam) and the unadapted individuals (John and Fred). For both symmetric and asymmetric distortions, it can be seen that transfer across identity was excellent. Indeed, after adaptation to Bill and Sam, the mean shift across subjects for John and Fred (symmetric = 5.37, asymmetric = 1.32) was, if anything, slightly larger than for Bill and Sam themselves (symmetric = 4.37, asymmetric = 0.75).

6.3.3 Experiment 10 - Discussion

Experiment 10 has revealed two new adaptation aftereffects for simple face shape distortions. Adapting to a face with both eyes up made an undistorted face appear to have its eyes shifted down, and vice versa. Adapting to a face with the left eye up and the right eye down made an undistorted face appear to have its left eye down and its right eye up, and vice versa.

The primary question in the present experiment was then whether some dimensions in face-space might be more adaptable than others. I contrasted two distortion types associated with different amounts of physical variability in the natural population of faces (symmetric = large variability; asymmetric = small variability), and confirmed that, in the unadapted state, subjects' perception followed this physical difference. That is, subjects tolerated a greater deviation in eye height before changing their rating of how much the face looked like the individual for symmetric distortions (both eyes up, or both eyes down) than for asymmetric distortions (left up right down, or right up left down).

When I then examined adaptation aftereffects, using an adaptor that was placed outside the threshold range for both distortion types, I found greater shift in the face rated as looking "most like Bill/Sam/John/Fred" for symmetric distortions than for asymmetric distortions (again, readers can test a similar result for themselves, using the ±50 pixel-per-eye stimuli in Figure 6.1 as adaptors). This finding suggests that (a) different dimensions of face-space can show different amounts of adaptation, and (b) for the moderate adaptor position tested in Experiment 10, dimensions which need to code a large range of stimulus values show more adaptation than dimensions which need to code a smaller range [7].

[7] Note again that the task was rating identity rather than normality. It is possible this lead to increased aftereffects for symmetric distortions compared to asymmetric ones as Cooper and Wojan (2000) suggest identity judgements are more affected by moving both eyes together whereas normality judgements are

Experiment 10 produced two other results of interest. First, adaptation took the form of a general shift in the curve; that is, the curve relating physical to perceived deviation remained the same shape, and shifted towards the adaptor direction. Second, adaptation generalised perfectly over individual identity. That is, despite the fact that the task deliberately referenced individual identity rather than generic normality, adaptation to Bill and Sam shifted the perception of John and Fred by as much as it shifted the perception of Bill and Sam. [8]

Both of these results are consistent with the theoretical idea illustrated in Figure 6.4, namely that adaptation produces a general shift in the norm of face-space along the adapted dimension that affects perception of all individuals. It does not support a view in which adaptation shifts the location of one individual person with respect to others. This general shift in the norm has not previously been demonstrated. With respect to adaptor position, previous studies have not tested distortions more extreme than the adaptor. With respect to face identity, Rhodes et al. (2003) showed that adaptation occurred with a change in identity but, in their study, the adaptor and test faces were always different individuals; thus the cross-identity adaptation was not compared to a same-identity condition. The Leopold et al. (2001) design also does not allow comparison of aftereffects for adapted and unadapted individuals; the trajectory along which adaptation takes place (e.g., Tim to anti-Tim) changes many aspects of the face simultaneously and thus contains only one individual (weaker and stronger versions of Tim). In contrast, by choosing simpler distortion types, multiple individuals can be changed in the same way along a given dimension, allowing comparison of the same distortion type for adapted and unadapted individuals.

---

affected by moving one eye. However, McKone et al. showed that when the amount of metric change in each distortion was matched (as was also the case here) Cooper and Wojan's link between symmetric distortion and identity and asymmetric distortions and normality did not hold. Further although there were slight differences between the identity and normality tasks in McKone et al, as noted below, results for symmetric and asymmetric were consistent across the two tasks.

[8] The generalisation of adaptation over identity should not be taken to suggest that subjects did not reference identity in this task. McKone et al. (submitted) used the same identity rating instructions as here, and contrasted these to normality ratings. In the unadapted state (adaptation was not tested), both tasks produced a Fechner law relationship between physical and perceived deviation. However, sensitivity to deviations was noticeably stronger in the identity task (i.e., smaller subthreshold region, and steeper slope-at-threshold). The present fit parameters agree well with those of McKone et al.'s identity task.

6.4 Experiment 11: Varying Adaptor Position

Experiment 11 had two aims. The primary aim was to test several adaptor positions, specifically ±5, ±20, and ±50 pixels-per-eye. These adaptor positions fall in different locations relative to the likely coding range for symmetric and asymmetric distortions of eye height.

For symmetric distortions, the coding range required to represent all possible eye heights is very large. Considering both the differences arising from individual face structure, and the even larger differences arising from three-dimensional up-down head rotation, the required range probably covers from the eyes being very close to the tip of the nose (looking down) to the eyes being very close to the hairline (looking up). In the current stimuli, these eye positions correspond to a range of approximately -60 to +60 pixels-per-eye. If this presumed range is correct, then, for symmetric distortions, all three adaptor positions will fall within the coding range of the relevant dimension in face-space. Under these circumstances, we might expect that all adaptor positions should produce strong aftereffects. The standard model of adaptation described in the General Discussion also predicts that the size of the aftereffect will become larger as the adaptor becomes more extreme.

The situation for asymmetric distortions might be very different. The required coding range for asymmetric distortions of eye height should be much smaller than for symmetric distortions. There is no three-dimensional head rotation that produces asymmetric changes in eye height, and the images in Figure 6.5 make it obvious that even a ±5 pixel-per-eye deviation would be unusual to observe in the natural population of faces. If we presume that the coding range for asymmetric distortions might be at most -10 to +10 pixels-per-eye, the ±5 adaptor could be within the coding range, while ±20 and ±50 adaptors would be outside it. It is therefore possible that an aftereffect with the ±5 adaptor might be stronger than the small effect identified for the ±20 adaptor in Experiment 10. Given that a ±50 adaptor is very far outside the coding range for asymmetric deviations, weak or no adaptation might be expected at this position.

In summary, the coding range hypothesis suggests that, as the adaptor is placed further from the pre-adaptation norm, the size of the aftereffect could increase for the symmetric distortion, but decrease for the asymmetric distortion. The difference between symmetric and asymmetric would then be greatest for ±50 adaptors, and smallest for ±5 adaptors. Whether any symmetric-asymmetric difference will remain with the ±5 adaptor cannot be predicted: this adaptor is probably within the required coding range for both distortion types, but it falls almost within the subthreshold region for symmetric (T=±4.89 pixels for N=19 in McKone et al., submitted) and well outside it for asymmetric (T=±1.79 pixels), and this may or may not affect the amount of adaptation.

The second aim of Experiment 11 was to replicate the results of Experiment 10 with a different measurement technique and more subjects. Instead of rating multiple deviation levels for how much each looked like Bill/Sam/John/Fred, an adjust-to-most-like procedure was used. Subjects were given an initially extreme distortion, and allowed to adjust the deviation level until the stimulus appeared most like the target individual. The effect of adaptation was again measured by comparing pre- and post-adaptation scores, with the specific measure (the shift in the stimulus chosen as most like Bill/Sam/John/Fred) being an equivalent of the shift-in-centre score reported in Experiment 10. The advantage of the adjust-to-most-like technique is that it is less time consuming to test one condition than the procedure used in Experiment 10, thus allowing more conditions to be tested. However, because subjects are no longer rating 27 deviation levels, the estimate of an individual subject's centre point is less reliable than in Experiment 10. The present experiment therefore used a multi-subject approach, where the intention was to present data only averaged over subjects.

6.4.1 Experiment 11 - Method

6.4.1.1 Subjects.

Twenty-four naïve subjects participated (age range 18-39 yrs; 9 male). Subjects were Caucasian members of the Australian National University community (mostly undergraduates), and were paid $25 for their participation (total 2 hrs). All reported normal or corrected-to-normal vision.

6.4.1.2 Design.

As in Experiment 10, all faces were upright. Each subject was tested in three 40 min sessions, with at least 24 hrs between sessions. Each session began with identity training on the original undistorted faces. Adaptation effects were then assessed for one adaptor position (e.g., -5 pixels-per-eye) for one distortion type (e.g., symmetric) and, after a 15 min filled delay, a different adaptor position (e.g., -50 pixels-per-eye) for the other distortion type (i.e., asymmetric). Across all sessions, a given subject received adaptors of only one direction (all positive or all negative), but was tested on all three adaptor levels (e.g., -5, -20, -50) for both distortion types. Order of the resulting six conditions was counterbalanced across subjects, within the constraints that each session contained one of each distortion type, and with different adaptor levels for each. Direction of adaptor (positive or negative) was manipulated between subjects, with the intention of collapsing over this variable, given that it had had no influence on adaptation effects in Experiment 10.

The final design was thus a 2 x 3 x 2 within-subjects design, crossing distortion type (symmetric, asymmetric) with adaptor level (5, 20, 50 pixels-per-eye) and identity status (adapted individuals = Bill and Sam; unadapted individuals = John and Fred). The subject's task was to adjust each face until it appeared "most like the original Bill/Sam/John/Fred". The physical deviation level the subject chose was recorded.

6.4.1.3 Stimuli.

For the pre- and post-adaptation phases, stimuli were as for Experiment 10, except that additional deviation levels were created. To allow adjustments in 1 pixel-per-eye increments, stimuli were prepared for every 1 pixel deviation between –40 and +40, for each identity and distortion type. These were prepared at the larger "test face" size used in Experiment 10. For the adaptation phase, additional stimuli were the ±5 and ±50 versions of Bill and Sam, prepared at the smaller "adaptor face" size.

6.4.1.4 Procedure.

The procedure was as for Experiment 10, except where noted.

Identity training phase. At the beginning of each session, four cycles of identity training were provided. Unlike in Experiment 10, training included only upright faces, shortening the training phase to approximately 10 mins. Each cycle first presented 8 trials of faces with names (2 each of Bill, John, Sam and Fred), followed by 16 trials of training-with-feedback without names (4 of each individual). One cycle through this identity training preceded testing of the second distortion type in the session.

At the beginning of Session 1, subjects were required to achieve 80% correct in the last of the four training cycles. Three subjects who initially failed to reach this criterion were given an extra two cycles before continuing. All subjects remained >80% correct in the training phase of Sessions 2 and 3.

Pre-adaptation phase. The pre-adaptation phase for a given condition (e.g., symmetric, -5 adaptor) presented two adjust-to-most-like trials for each of the four individuals. One of these trials began with an obvious positive distortion (e.g., eyes too high), and one with an obvious negative distortion (e.g., eyes too low). The starting position was chosen randomly for each trial within the range +20 to +40 pixels-per-eye, or -20 to -40 pixels-per-eye, respectively.

To adjust the stimuli, subjects used number keys at the top of the keyboard. Keys 1 and 4 allowed coarse adjustment, moving the eyes by ±5 pixels with each press (one key for each direction). Keys 2 and 3 allowed finer adjustments, moving the eyes by ±1 pixel. A beep sounded if subjects tried to move the eyes outside the maximum range of ± 40. Subjects were instructed to make adjustments fairly quickly, taking approximately 10 key presses, and at most 1 s per press. When the subject considered the stimulus to "look like the undistorted person", they pressed the space bar to accept this stimulus, and its deviation level was recorded.

Adaptation phase. This was exactly as for Experiment 10.

Post-adaptation phase. This was as for the pre-adaptation phase with the addition of 5 sec top-up adaptors, shown in the same format as in Experiment 10.

Delays within and across sessions. The extensive (24 hour) delay between sessions was necessary to avoid carryover of adaptation between successive tests of the same distortion type. In initial pilot testing (N=6), I tried including all adaptor levels for a given distortion type in a single session. However, there was noticeable carryover from one adaptation condition to the next "pre"-adaptation measurement, which occurred despite a 3 min intervening task showing 60 novel undistorted individuals, and a 2 min cycle of identity training. A second round of pilot testing (N=5) increased the first intervening task to 13 mins. This presented 240 trials of greyscale pictures, each for 3 s. The stimulus for each trial was chosen at random from 81 novel individual faces and 51 other objects (at least 120 trials were faces). To ensure attention, subjects were told to press the space bar each time one of the pictures appeared for a second or successive time. Even with this lengthier filler task, some pilot subjects revealed substantial carryover, and the mean effect was 3.08 pixels (SEM = 2.29).

In the final procedure, the delay between testing two adaptor position conditions for the same distortion type (e.g., -5 symmetric, and -50 symmetric) was a minimum of 24 hrs. Under these conditions, there was no carryover. This is demonstrated in Table 6.1, which presents pre-adaptation results for the first, second and third session. For a given distortion type, any carryover would be revealed as a cumulative change in the stimulus chosen as most like the individual, such that stimuli would become more negative for subjects who received negative adaptors (N=12), and more positive for subjects who received positive adaptors (N=12). This did not happen. Linear trend analysis across sessions (1 – 3), conducted for each distortion type and direction separately, revealed no significant trends, all *p*s > .3.

Table 6.1. Assessing carry-over effects across the three sessions in the pre-adaptation tests: Mean deviation of the stimulus chosen as most like Bill/Sam/John/Fred (and SEM) for each distortion type.

	Session 1	Session 2	Session 3
Asymmetric			
Positive adaptors in all sessions	0.17 (0.14)	0.19 (0.15)	0.26 (0.16)
Negative adaptors in all sessions	0.32 (0.24)	0.22 (0.24)	0.49 (0.14)
Symmetric			
Positive adaptors in all sessions	-1.48 (0.88)	-1.09 (1.05)	-0.56 (0.81)
Negative adaptors in all sessions	-2.55 (0.61)	-1.36 (0.63)	-1.86 (1.07)

With respect to testing different distortion types, asymmetric and symmetric were included within the same session, with a total of 15 mins filled delay between them. This comprised the 13 min filler task described above, plus the 2 mins for one cycle of identity training. Note that any adaptation effect left over from one distortion type could not affect adjustment performance on the other; for example, adaptation to an asymmetric distortion could not be revealed on a task that allowed only symmetric adjustments.

6.4.2 Experiment 11 – Results

Trials on which subjects made no adjustments before pressing the spacebar were treated as errors and discarded. This never occurred for more than one adjustment trial for any given subject.

6.4.2.1 Confirming greater perceived variability for symmetric than asymmetric distortions.

As for the previous experiment, difference in perceived variability between the symmetric and asymmetric distortions was assessed in the unadapted state. For each subject, the pre-adaptation deviation level (measured in pixels-per-eye) of the stimulus chosen as "most like Bill/Sam/John/Fred" was computed for each distortion type. To give the most reliable data, this was done averaging over all four individuals, and all three future adaptor levels (e.g., -5, -20, and -50).

The variability of the symmetric and asymmetric conditions across subjects was then compared using Levene's test of equality of variances. This showed that the variance of 5.38 for the symmetric distortion (around a mean of –1.49) was significantly higher than the variance of 0.19 for the asymmetric distortion (around a mean of 0.27), $F(1, 46) = 30.26, p < .001$. That is, subjects were less certain about where the eyes should be placed in the symmetric condition, corresponding to the finding that there was a broader subthreshold region for ratings of the symmetric distortion than the asymmetric distortion in Experiment 10. Again, the implication is that the current subjects' perception tracked the fact that the symmetric distortion corresponds to more variability in the real world than does the asymmetric distortion.

6.4.2.2 Adaptation effects.

As in Experiment 10, shift scores were defined such that a positive shift always corresponded to a change in the direction predicted for adaptation. If an adaptation aftereffect is present then, after adapting to an eyes up face (for example), eyes up stimuli will be perceived as more like "Bill" than previously, while eyes down stimuli will be perceived as less like "Bill". Thus, after adaptation to positive adaptors, the stimulus chosen as "most like" should be more positive than before adaptation (thus, *shift* = $c_{post}–c_{pre}$), and after adaptation to negative adaptors, the chosen stimulus should become more negative (thus, *shift* = $c_{pre}–c_{post}$).

An initial 4-way mixed ANOVA was conducted on the shift scores. This included distortion type (symmetric vs. asymmetric), adapted/unadapted identity status (Bill and Sam vs. John and Fred), adaptor level (5, 20, 50), and direction of adaptor (positive or negative). This showed no main effect of direction or any interactions involving this variable (direction by distortion type, $F(1, 22) = 3.27, p > .08$, all other Fs < 1). Thus, I collapsed over direction in later analyses.

Generalisation over identity. The 4-way ANOVA revealed no main effect of identity status, $F(1, 22) = 2.88, p > .1$. However there was a significant interaction between identity status and distortion type, $F(1, 22) = 5.97, p < .05$. For the asymmetric distortion, results agreed with those of Experiment 10. Adaptation generalised perfectly over identity: if anything, unadapted individuals showed a slightly larger effect (John and Fred, M = 0.78, SEM=0.12) than unadapted individuals (Bill and Sam, M= 0.61, SEM=0.15). For the symmetric condition, however, results differed from those of the first experiment: adaptation transferred well, but not perfectly, over identity changes. The effect for unadapted individuals (John and Fred, M = 2.29, SEM= 0.48) was significantly smaller than for adapted individuals (Bill and Sam, M = 3.24, SEM = 0.68), $t(23) = 2.22, p < .05$. Note that these results are averaged across level of adaptor, because there was no 3-way interaction between identity status, distortion type, and adaptor level, $F < 1$. A check of the three adaptor levels separately confirmed that the same trend was present for symmetric distortions at all adaptor levels (for ±5, J&F=1.25, B&S=1.75; for ±20, J&F=2.97, B&S=3.66; for ±50, J&F=2.65, B&S=4.32).

Effects of distortion type and adaptor level. Given the lack of any 3- or 4-way interactions in the initial ANOVA, the questions of primary interest – how the adaptation aftereffect was influenced by distortion type and adaptor position – were examined collapsing over identity and direction. Results are shown in Figure 6.11. A 2-way ANOVA revealed a main effect of distortion type, $F(1, 23) = 12.64, p < .01$, and a significant interaction between distortion type and adaptor position, Wilks' lambda = .70 [9], $F(1, 23) = 4.67, p < .05$. As can be seen in Figure 6.11, the interaction takes the form that, while shift scores increased with increasing adaptor level for the symmetric

[9] Where the sphericity assumption of repeated measures ANOVA was violated, the multivatiate approach was used.

distortion (the trend was 5<20<50), the opposite was true for the asymmetric distortion (the trend was 5>20>50).

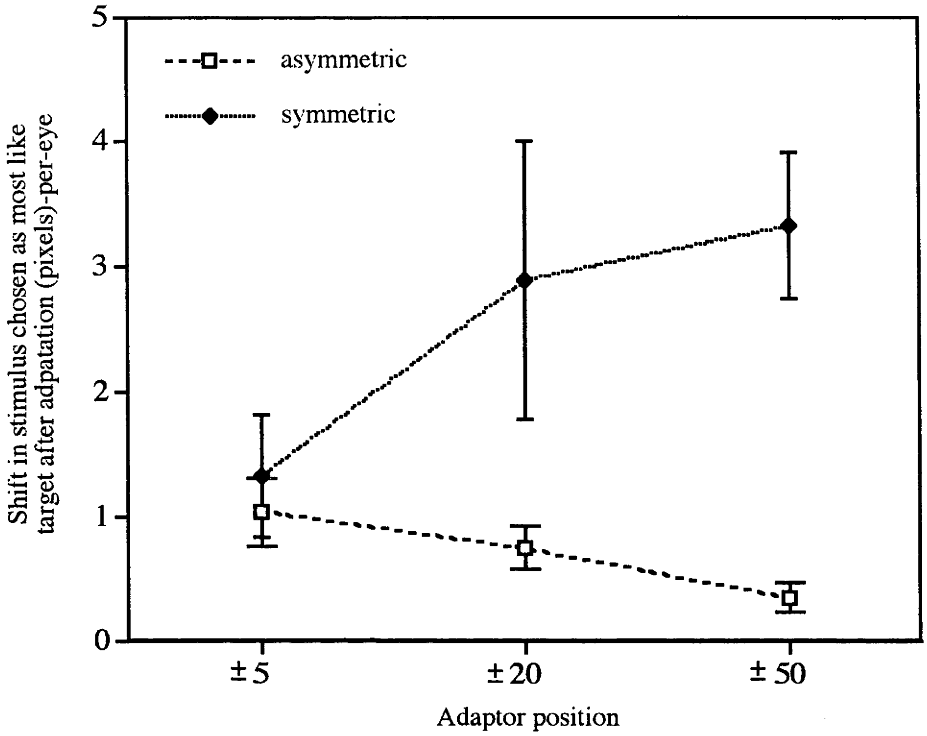


Figure 6.11. Experiment 11: Interaction between distortion type (symmetric vs. asymmetric) and adaptor position (±5, ±20, ±50). Data are collapsed over identity and direction of adaptor. Error bars show ± 1 SEM.

To follow up the interaction statistically, results for the two distortion types were analysed separately. Results of the analysis were neatest for the asymmetric distortion, so this is presented first. In this condition, trend analysis revealed a significant linear trend, $F(1, 23) = 5.82, p < .05$, with no quadratic component, $F < 1$. Contrasts confirmed that the ±5 pixel adaptor produced a significantly larger aftereffect than the ±50 pixel adaptor, $F(1, 23) = 6.65, p < .05$. The ±20 pixel condition fell in the middle, and was not significantly different from either the ±5 condition, $F(1, 23) = 1.21$, $p > .05$, or the ±50 condition, $F(1, 23) = 2.18, p > .05$.

For the symmetric distortion, there was a trend in the opposite direction (see Figure 6.11). The sphericity assumption of repeated measures ANOVA was violated so trend analysis was not possible. The multivariate equivalent of a 1-way ANOVA was not significant, Wilks' lambda = .80, $F(1, 22) = 2.73, p = .087$. However, this test does not take account of the ordering of the conditions, and thus has substantially less power than trend analysis. *A priori* t-tests showed that the ±5 adaptor produced a significantly smaller aftereffect than the ±50 adaptor, $t(23) = 2.36, p < .05$. The ±20 adaptor again fell in the middle, and did not differ significantly from either the ±5 adaptor, $t(23) = 1.31, p > .05$, or the ±50 adaptor, $t < 1$. Overall, these results confirm that, for symmetric distortions, adaptors with more extreme deviations produced more adaptation, and for asymmetric distortions, they produced less adaptation.

I also compared amounts of adaptation for symmetric and asymmetric at each adaptor level. For the ±50 adaptors (i.e., within the coding range for symmetric, but very far outside it for asymmetric) the greater adaptation for symmetric than asymmetric distortions was significant, $t(23) = 4.90, p < .001$. The difference in the same direction for ±20 adaptors was not quite significant, $t(23) = 1.97, p = .06$, although it can be seen from Figure 6.11 that this was probably attributable to the abnormally large variability in the symmetric ±20 condition [10]. For ±5 adaptors (i.e., within the coding range for both distortions), a quite different result was revealed. Here, symmetric showed only marginally more adaptation than asymmetric, and this difference was far from significant, $t<1$. Overall, these results replicate the symmetric-asymmetric difference found in Experiment 10 with moderate and extreme adaptors, but also show that the effect is significantly modified by adaptor position (i.e. coding range).

Finally, adaptation was compared to zero in each condition in Figure 6.11. For symmetric distortions, one-sample t-tests showed significant adaptation for all three adaptor levels (all *p*s < .02). For asymmetric distortions, the generally smaller effects were still all significant (*p*s < .01), including, somewhat surprisingly, the tiny (0.34 pixels) effect at ±50.

6.4.3 Experiment 11- Discussion

The most interesting finding of Experiment 11 was of the very different influence of adaptor position on the aftereffect for the two distortion types. For symmetric distortions, the more extreme the adaptor, the larger the adaptation effect.

[10] Note that this larger variability may be a real finding. There are likely to be differences between subjects in the width of the threshold region (corresponding to slight differences in the shape and/or steepness of curves presented in the model in the General Discussion). This would then change the exact position of ±20 pixels relative to the threshold for individual subjects.

For asymmetric distortions, the more extreme the adaptor, the smaller the adaptation effect.

The pattern found for symmetric can be interpreted in terms of a standard neural model of norm-based aftereffects derived from low- and mid-level vision, in which the size of the aftereffect increases as the adaptor shifts further from the norm but remains within the coding range of the relevant dimension. As will be discussed in detail in the General Discussion, the pattern found for asymmetric is partially, but not entirely, consistent with this model. Briefly, the fact that adaptation was largest with the ±5 adaptor is consistent with the idea that strong adaptation occurs only when the adaptor is placed within the narrow coding range for asymmetric distortions. However, the finding of weaker adaptation at ±20 and ±50 argues that these adaptor positions must have fallen outside the coding range and, if this is the case, the standard model predicts <u>no</u> aftereffect, rather than merely a smaller but nonzero effect.

With respect to identity, Experiment 11 showed good generalisation of adaptation over adapted/unadapted identity status. This is consistent with Rhodes et al. (2003) and the present Experiment 10. It supports the idea that adaptation causes a general shift in a face-space norm (Figure 6.4), in that adapting to Bill and Sam not only changes perception of these particular individuals, but also changes perception of John and Fred. Generalisation results were not exactly as for the first experiment. Previously, there was perfect generalisation over identity for both symmetric and asymmetric; indeed, in both cases, unadapted individuals showed slightly <u>more</u> adaptation than adapted individuals. In the present experiment, there was again perfect generalisation for asymmetric distortions, but less than complete generalisation for symmetric distortions. From a theoretical perspective, less than complete generalisation is not necessarily a surprise, in that the adjust-to-most-like task referenced individual identity. However, the rating task of Experiment 10 also referenced individual identity. Thus, the reason for the difference in symmetric results across experiments is not clear. My only suggestion is that some subjects show greater identity generalisation than others (and it can be noted that Experiment 10 tested only 5 participants, while Experiment 11 tested 24).

A final point is that the overall effect of adaptation in Experiment 11 was somewhat smaller than in Experiment 10. Comparing the data for the adaptor level that appeared in both experiments (±20), the mean shift in Experiment 11 was 3.07 pixels for symmetric and 0.74 pixels for asymmetric, and the mean shift in Experiment 10 was

4.98 pixels for symmetric and 1.44 pixels for asymmetric. I presume that the smaller effects in the "adjust-to-most-like" task than in the rating task can be attributed to the time required to make an adjustment (on average, subjects made approximately 8 - 11 adjustments taking 7 – 16 s per trial), reflecting some decay of adaptation in the period immediately after removal of each top-up adaptor (e.g., see Leopold et al., 2001), or the lengthened presentation of the test stimulus (Leopold, Rhodes, Muller, & Jeffery, in press).

6.5 Experiment 12: Transfer Of Adaptation Between Orientations

The aim of Experiment 12 was to show that the adaptation found in Experiments 10 and 11 for upright faces had its origin in the face recognition system, by demonstrating zero transfer of adaptation between upright adaptors and inverted test faces (and vice versa). Previous research has established that adaptation occurs for inverted faces with inverted adaptors (I-I condition); indeed, it is commonly as strong as for upright faces with upright adaptors (Leopold et al., 2001; Watson & Clifford, 2003; Webster & MacLin, 1999). This does not necessarily mean that adaptation in the U-U and I-I conditions is coming from the same source. For example, adaptation for inverted faces could arise from generic object recognition processes and/or from shape representations in mid-level vision (Kourtzi & Kanwisher, 2001; Regan & Hamstra, 1992), rather than from face-specific mechanisms.

To test for different sources, examining transfer across orientations is the relevant technique. In previous studies, quite strong transfer has been the usual finding. Specifically, an upright adaptor produces a strong perceptual aftereffect on an inverted test face and, although the transfer from an inverted adaptor to an upright test face is usually weaker, even this is greater than zero (Watson & Clifford, 2003; Webster & MacLin, 1999). These findings indicate that, in previous studies, adaptation for upright faces and adaptation for inverted faces originated from neural populations that were at least partly overlapping, rather than distinct.

Experiment 12 tested the idea that, with the relational-change distortion types used in current experiments, adaptation in the orientation mismatch conditions might be reduced to zero, thus indicating that adaptation for upright faces could be attributed to face-specific sources. Both symmetric and asymmetric distortions were tested for three orientation conditions, namely inverted adaptor-inverted test (I-I), upright adaptor–inverted test (U-I) and inverted adaptor-upright test (I-U). The experiment used the full-curve rating task of Experiment 10, and the same adaptor position (±20 pixels). The rating method was used in preference to the adjustment procedure because it produced larger adaptation effects, making any conclusion of no adaptation in the orientation mismatch conditions more convincing.

6.5.1 Experiment 12 - Method

6.5.1.1 Subjects.

The experiment used a single-subject approach. Subjects were the four naïve individuals described in Experiment 10 (S1-S4), paid $10 per session.

6.5.1.2 Design, Stimuli & Procedure.

In addition to the four sessions of Experiment 10 (U-U), each subject participated in a further 12 sessions of approximately 40 mins each (8 hrs for 12 sessions). Each session tested a single adaptor-test orientation condition (e.g., U-I) for one distortion type (e.g., symmetric), and one adaptor direction (+20 or -20). For each subject, all conditions for I-I and U-U (from Experiment 10) were tested first; these were interleaved and, as far as possible, tested in counterbalanced order across subjects. After this, all conditions for I-U and U-I were tested; these were again interleaved and counterbalanced as far as possible. The minimum delay between sessions was 30 mins (and the maximum 18 days); the delay between two sessions testing the same distortion type was 1 hr in one case (between U-I –20 symmetric and I-U +20 symmetric, for S4), and more than 12 hrs in all other cases, suggesting that carryover should not have been a problem.

Experiment 12 used the same task as Experiment 10 (i.e., rating 27 deviation levels for how much each looked like Bill/Sam/John/Fred). Stimuli and procedure were as for Experiment 10, with the exception of the changes in stimulus orientation. Inverted versions of all stimuli were created by rotating the upright versions by 180°. Note that the identity training at the beginning of each session included both upright and inverted orientations. Also note that, for a given condition, orientation for test faces in the pre- and post-adaptation phase was always the same. As an example, consider the inverted adaptor-upright test condition. Here, the pre-adaptation phase included only upright test faces for rating, the adaptor phase showed inverted faces, and the post-adaptation phase included upright test faces for rating with each trial preceded by an inverted top-up adaptor.

6.5.2 Experiment 12 - Results

6.5.2.1 Inverted tests faces and inverted adaptors (I-I).

Before turning to the transfer conditions, I first examined aftereffects when the adaptor and test faces were both inverted. In the unadapted state, the open circles in Figure 6.12a show identity ratings against deviation level for inverted faces, averaged over the four subjects. Data are collapsed across identity and direction of future adaptor. The figure shows that Fechner's law still provided good fits to the data for inverted test faces.

The pre-adaptation ratings also show that sensitivity to distortion was weaker in inverted faces than in upright faces (compare Figure 6.12a to Figure 6.8). The subthreshold region was wider in the inverted than in the upright orientation, for both symmetric distortions (upright = 11.62, inverted = 14.83) and asymmetric distortions (upright = 8.49, inverted = 9.89). These findings are consistent with extensive previous evidence that relational changes (altering spacing between features while keeping local feature shape constant) affect perception of upright faces more strongly than perception of inverted faces (e.g., Bartlett & Searcy, 1993; Gilchrist & McKone, 2003; Murray, Yong, & Rhodes, 2000).

Finally, with respect to adaptation, Figure 6.12a shows that the aftereffect took the form of a simple shift in the rating curve in the direction of the adaptor. This allowed shift to be calculated as for upright faces in Experiment 10 (for positive adaptors, $shift = c_{post}-c_{pre}$; for negative adaptors, $shift = c_{pre}-c_{post}$). Figure 6.12b then shows shift scores for individual subjects. As for upright faces, the aftereffect was larger for symmetric than for asymmetric distortions, and this again corresponded to a wider threshold in the pre-adaptation ratings. Generalisation over identity for I-I was complete, for both symmetric distortions (for adapted identities, Bill and Sam = 5.00; for unadapted identities, John and Fred = 5.11), and for asymmetric distortions (Bill and Sam = 1.76; John and Fred = 2.51).

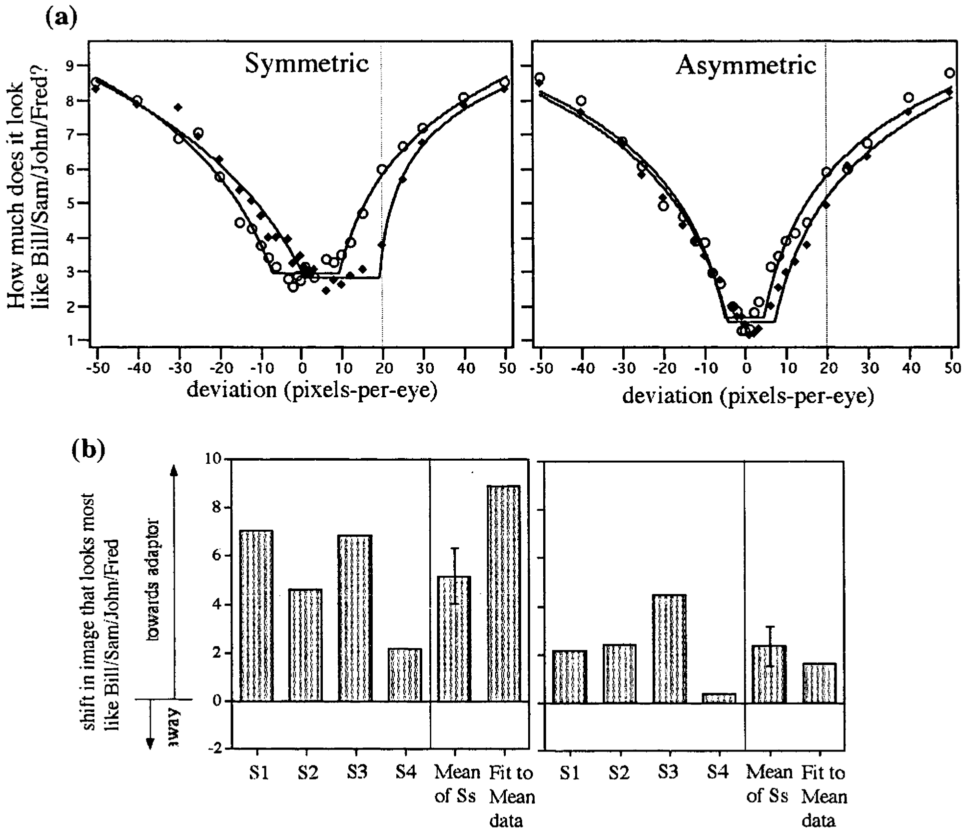


Figure 6.12. Experiment 12: (a) adapted (filled diamonds) and unadapted (open circles) data and fits for the inverted adaptor to inverted test face condition (I-I) averaged over subjects. To present results efficiently, the data are shown collapsed over identity status (adapted vs. unadapted individuals) and also direction of adaptor: results for negative adaptor sessions have had the x-axis flipped (for both pre-adaptation and post-adaptation) before averaging with positive adaptor sessions, so that the adaptor is notionally at "+20" (shown as a dotted line). Also shown (b) are the centre-shift scores for individual subjects (as described in Figure 6.10, but collapsed over identity of adaptor).

6.5.2.2 Transfer of adaptation across orientation.

Turning to transfer across orientations, the major result of Experiment 12 is shown in Figure 6.13 as the shift score for each subject. Essentially no adaptation occurred in either of the two orientation mismatch conditions (I-U and U-I). Of the 16 measurements in the individual-subject data, only one (S3 for U-I symmetric) shows any noticeable adaptation. For the fit of the subject-averaged data, the adaptation effect was tiny for both symmetric distortions (U-I = –0.08 pixels, I-U = 0.29) and asymmetric distortions (U-I = 0.09, I-U = 0.64). The lack of transfer for symmetric distortions is particularly compelling given the substantial effects for this distortion type for both U-U (Figure 6.10) and I-I (Figure 6.12).

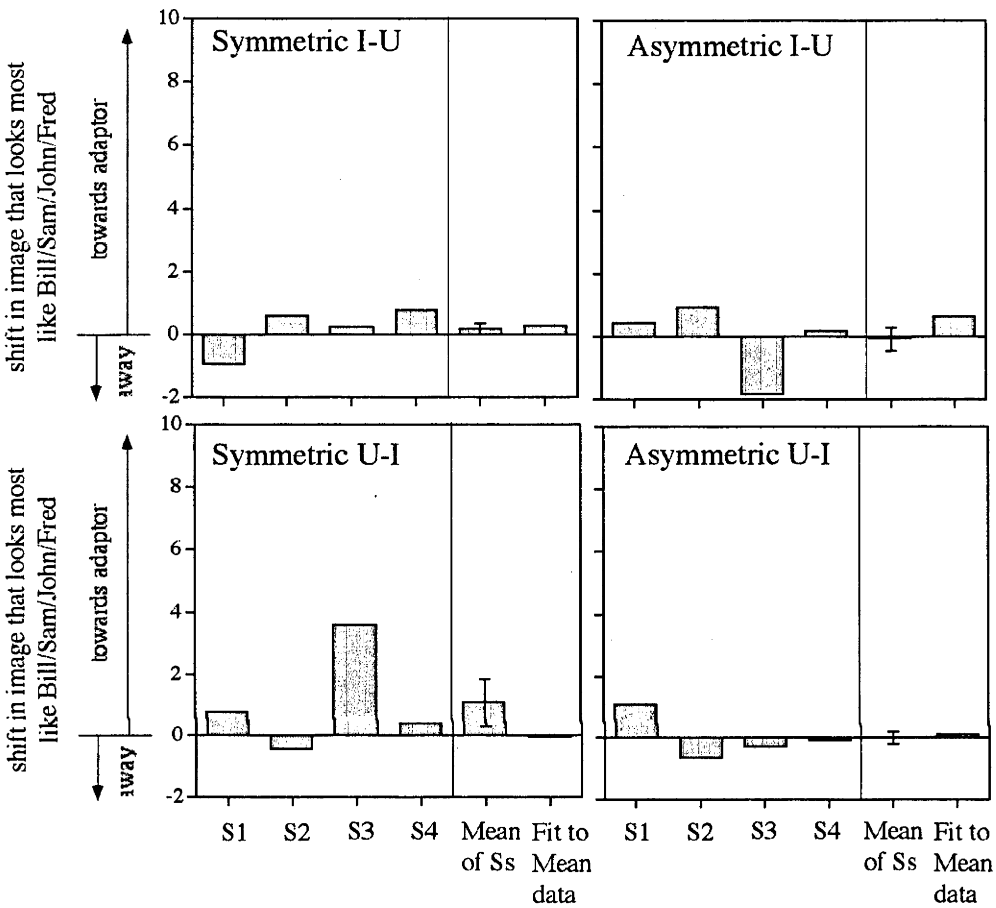


Figure 6.13. Experiment 12: Effect of mismatching orientation on centre-shift score, showing lack of transfer from either upright adaptors to inverted test faces (U-I) or from inverted to upright (I-U). Note that the scale is the same as for figures showing U-U (Figure 6.10) and I-I (Figure 6.12). Data are collapsed over identity and direction of adaptor.

6.5.3 Experiment 12 - Discussion

The finding of strong adaptation when both adaptor and test faces are inverted (I-I) replicates several previous findings with other distortion types (Leopold et al., 2001; Watson & Clifford, 2003; Webster & MacLin, 1999). It is important to note that this adaptation does not necessarily come from the same source as that for upright faces.

Substantial evidence argues that inverted faces are not processed by the same system as upright faces (e.g., Rhodes, Jeffery et al., 2004; Tanaka & Farah, 1993; Young et al., 1987). Some coding of information about distances between parts must take place within non-face object recognition system/s (even if this coding is coarser than takes place within the face-specific system), and adaptation may arise within these system/s for inverted faces.

The results of the current orientation-mismatch conditions argue that this was indeed the case. Experiment 12 obtained essentially no transfer of adaptation across orientations (see Figure 6.13). This demonstrates a reliance on distinct, rather than overlapping, neural populations for upright and inverted faces. Thus, the aftereffects for upright faces in Experiments 10 and 11 must have come from adaptation of cells that specifically process upright faces, rather than from other sources of adaptation.

I presume that the use of distortion types that are specifically relational in nature was important in achieving this result. (Another relevant factor might have been tasks that referenced individual identity, rather than merely face normality.) The global expansion/contraction distortions used in previous studies change local feature shape as well as the spacing between features (Rhodes et al., 2003; Leopold et al., 2001; Watson & Clifford, 2003; Webster & MacLin, 1999). With local feature distortions, a substantial component of adaptation could come from mid-level shape coding mechanisms that would show strong transfer across inversion. For example, widened eyes could adapt a direction-of-elongation mechanism (Regan & Hamstra, 1992) to make the original eyes appear narrowed, and such adaptation would be unaffected by 180° rotations.

6.6 General Discussion

The major findings of the present study were that (a) for upright faces, symmetric distortions of eye height produced more adaptation than asymmetric distortions when the adaptor was placed well away from the original norm; (b) adaptation for symmetric distortions increased as the distance of the adaptor from the norm became more extreme (±5 vs ±20 vs ±50 pixel-per-eye deviation), whereas adaptation for asymmetric distortions decreased (although even for ±50, it remained significantly above zero); and (c) adaptation disappeared when adapting and test faces differed in orientation, indicating that the adaptation for upright faces arose from cells tuned specifically to the upright orientation. Other findings were: (d) there was good or perfect generalisation of adaptation over identity; and (e) the form of adaptation was a simple shift in the curve relating physical to perceived distortion.

I now consider the theoretical implications of these results. I deal first with the primary results, namely those for upright faces, in terms of understanding how individual faces are coded in face-space, and in terms of developing neural models of adaptation for face-space dimensions. Following this, I briefly consider how adaptation for inverted faces is to be explained.

6.6.1 Some dimensions in face-space are more adaptable than others.

For upright faces, the representation of individual identity has commonly been understood in terms of face-space. The current results regarding the form of adaptation, and regarding the generalisation over identity, are consistent with the idea that adaptation shifts the face-space norm (Leopold, et al., 2001; Rhodes et al., 2003). This shift can be thought of as occurring towards the adaptor, in a trajectory along the adapted dimension/s, as illustrated in Figure 6.4.

My major result is then that adaptation differed for symmetric and asymmetric eye-height distortions. Previous researchers (Leopold et al., 2001; Rhodes et al., 2004, in Rhodes, Robbins, et al., in press) have demonstrated adaptation effects that are interpretable within the context of face-space, but have not attempted to contrast adaptation for different dimensions. The present study provides the first evidence showing that the norm of face-space can be shifted more easily in some directions than in others. Specifically, I found that one dimension in face-space (symmetric eye height) is more adaptable than another (asymmetric eye height), both in the sense of showing a larger aftereffect for certain fixed adaptor positions, and in the sense of showing strong adaptation to a wider range of adaptor positions.

In explaining the specific origin of the symmetric-asymmetric difference, I have argued that dimensions which need to code a larger range of stimulus values are more adaptable than those which need to code a smaller range. Symmetric deviations of eye height are associated with large variability in the range of natural face images, partly due to large physical variability across individuals (Farkas et al., 1994), and partly due to up-down head rotation (e.g., nodding) causing further large changes in apparent eyes-hairline and eyes-chin distance. Asymmetric differences in eye height are associated with much smaller variability: physical measurements indicate little variance in asymmetry between the eyes across individuals (Hreczko & Farkas, 1994), and there is no three-dimensional head rotation which can alter eye asymmetry with respect to the main axis of the face.

6.6.2 What frame of reference does face-space use for coding eye position?

Before turning to possible neural models of the symmetric-asymmetric differences, it is worth noting that the results demonstrate something important about the reference point/s that face-space uses to code eye height. In particular, they indicate that eye height is not coded in terms of individual eye position, nor is it coded with respect to only one half of the face (i.e., left half or right half). This conclusion can be drawn because the symmetric and asymmetric distortions matched total amount of metric deviation from the original. Considering the right eye, for example, this means that a symmetric distortion of +20 pixels-per-eye shifted the right eye up by 20 pixels, and an asymmetric distortion of +20 pixels-per-eye also shifted the right eye up by 20 pixels. Thus, if what was happening on the other side of the face was irrelevant, then the amount of adaptation for symmetric distortions (right eye up, left eye up) should have been the same as the amount of adaptation for asymmetric distortions (right eye up, left eye down). The fact that this did not happen indicates that eye height is not coded independently for each eye; instead, the position of each eye must be coded in some way that takes into account the position of the other.

The present results do not indicate whether the frame of reference for coding eye position is with respect to distance from the other features or with respect to the head outline (e.g., the hairline and chin). Researchers who have investigated sensitivity to relational changes have argued that distance to other features (e.g., eyes-nose distance) is important. However, the very large aftereffect obtained with a symmetric ±50 adaptor suggests that the height-of-both-eyes-together might also be coded with respect to face outline: it is only by allowing for up-down head rotations (which place the eyes near the hairline) that there would be any need to code a range of eye heights as extreme as 50 pixels.

6.6.3 Neural models of adaptation for face-space dimensions.

At several places in this chapter, I have referred to predictions derived from a specific neural model of adaptation, which can be applied to dimensions of face-space. I now present this model in detail, and demonstrate how it is able to explain most, but not all, of my findings.

Face adaptation aftereffects occur with respect to a single perceptual centre point (in the present study, a face with average eye height properties). In low- and mid-level vision, aftereffects of this kind have been successfully explained using models (see Figure 6.14) based on three simple assumptions. First, the coding of the dimension in question relies on two broadly tuned mechanisms that respond maximally to stimuli on opposite sides of the norm. For example, Regan and Hamstra (1992) explained coding of aspect ratio by presuming that, at some level in the visual pathway, neurons are organised functionally into two pools, one of which is preferentially excited by vertically elongated stimuli (Pool 1), and the other of which is preferentially excited by horizontally elongated stimuli (Pool 2).

Second, any given stimulus produces activation in both pools but, unless the stimulus is at the norm, this activation will be stronger in one pool than the other. Perception is then determined by considering the output from both pools simultaneously (as either the difference or the ratio of the outputs). Thus, for example, a vertically elongated stimulus is perceived as "tall and thin" because Pool 1 is firing more than Pool 2; a square is perceived as square because Pool 1 and Pool 2 are firing equally; and a horizontally elongated stimulus is perceived as "short and fat" because Pool 2 is firing more than Pool 1.

The third basic assumption is that adaptation causes each pool to reduce its firing rate in proportion to the strength of its initial unadapted response (e.g., Maddess, McCourt, Blakeslee, & Cunningham, 1988). After adaptation to a vertically elongated stimulus, therefore, the response of Pool 1 will be reduced more than the response of Pool 2, shifting the point of equal response levels – and the physical stimulus now perceived as being square – in the direction of vertical elongation.

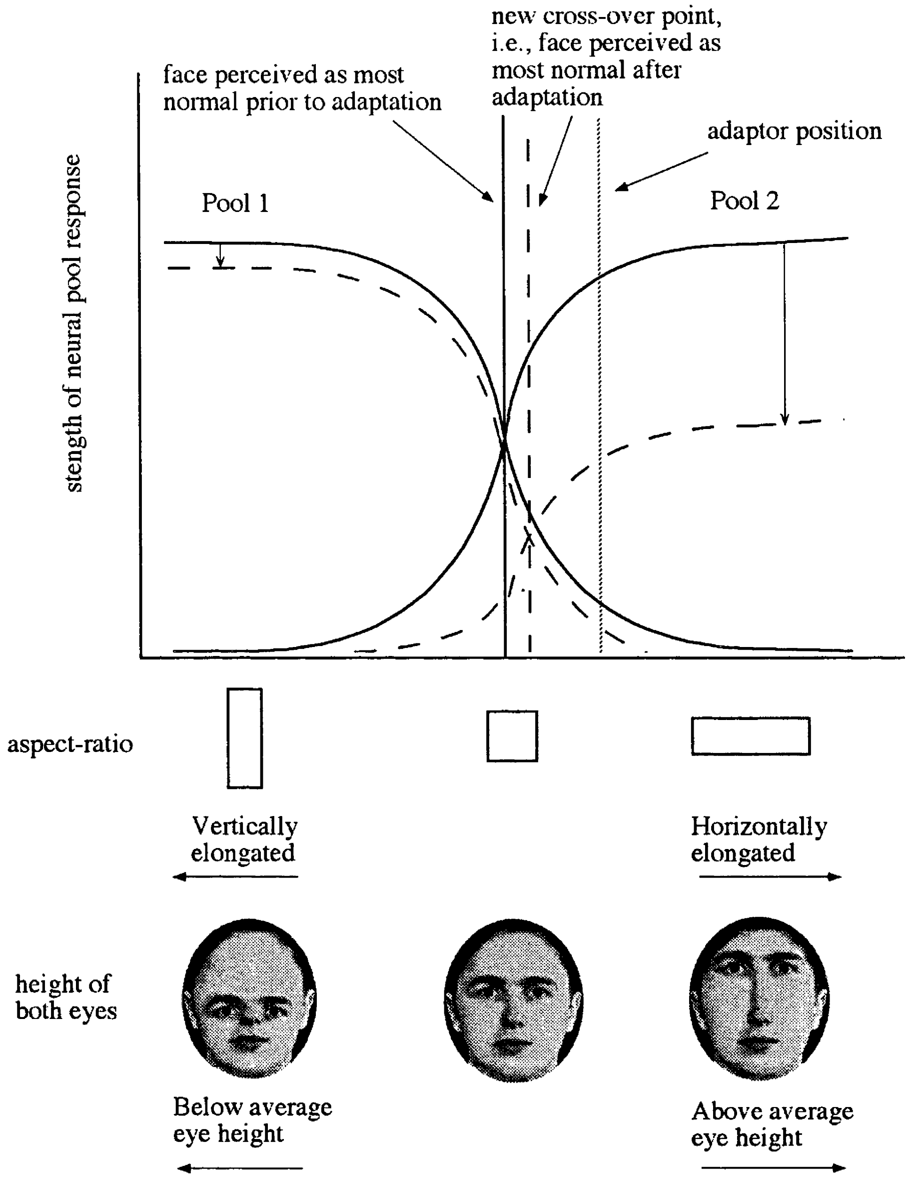


Figure 6.14. Two-pool neural mechanism for coding an opponent dimension, with an explanation of how the adaptation aftereffect occurs. Two pools of cells code deviations on each side of the norm. The cross-over point of these then gives the stimulus perceived as most normal. Solid lines show the response of pools of cells without adaptation. Dashed lines show the response after adaptation, that is, the perceived norm shifts towards the adaptor because Pool 2 is reduced by a larger amount than Pool 1.

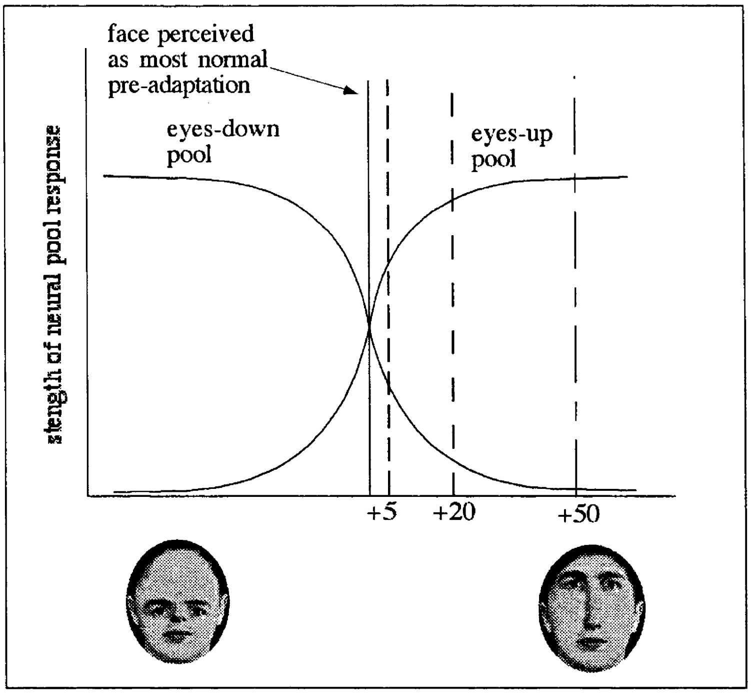
I now consider how well this type of model can explain the current face aftereffects. Figure 6.14 indicates how the model can be applied to coding of a height-of-both-eyes-together dimension in face-space (corresponding to the symmetric distortion), where the two pools are preferentially excited by eyes-up and eyes-down stimuli respectively. In the unadapted state, an undistorted face stimulus is perceived as most like the target individual because the eyes-up pool and the eyes-down pool are firing equally. After adaptation to an eyes-up face, the response of the eyes-up pool will be reduced more than that of the eyes-down pool, shifting the stimulus perceived as most like the target (i.e., the crossover point of equal response from both pools) towards the eyes-up direction, and making the original undistorted stimulus perceived as having its eyes down. This idea also implies that perception on both sides of the norm will shift in same direction (i.e., both an eyes-down stimulus <u>and</u> an eyes-up stimulus will appear to have higher eyes after adaptation than before). This is consistent with the current results indicating a simple shift in the overall rating curve.

Turning to the effects of adaptor position, the model predicts that the size of the aftereffect will increase as the adaptor becomes more extreme within the range covered by the two pools. This is because the <u>relative</u> change in the output of the two pools – and thus the amount of shift in the crossover point – will be greatest when the adaptor is located in a region producing very different initial responses from each pool (e.g., very strong response from Pool 1, very weak from Pool 2). This is illustrated in Figure 6.15a. For symmetric deviations of eye height, this figure assumes a large coding range of -60 to + 60 pixels-per-eye (see Introduction to Experiment 11 for justification), and marks the approximate location of the +5, +20 and +50 adaptors within this range. For a 5 pixel eyes-up (+5) adaptor, the shift in the norm after adaptation will be relatively small, because the eyes-up and eyes-down pools are initially firing at only moderately different levels, meaning that the reduction in firing rate for the eyes-up pool will be only a small amount more than for the eyes-down pool. For a +20 adaptor, the shift in norm will be larger, because the initial difference in firing rate between the two pools is larger. For a +50 adaptor, the initial difference is larger again and, indeed, the firing of the eyes-down pool would barely be affected at all while the eyes-up pool would be strongly affected, leading to the largest possible shift in the norm [11]. In the case of the symmetric distortion, these predictions of the "large range two-pool model" (Figure 6.15a) fully match the results, in which the aftereffect increased with increasing adaptor distortion.

---

[11] For adaptors beyond ±60, the normal "first-order" configuration of the face (i.e., eyes <u>below</u> hairline and <u>above</u> nose) would be broken. I suspect that adaptors outside this range (e.g. an adaptor with the eyes above the top of the head) would probably not produce an adaptation aftereffect.

(a) Large range two-pool model



(b) Small range two-pool model

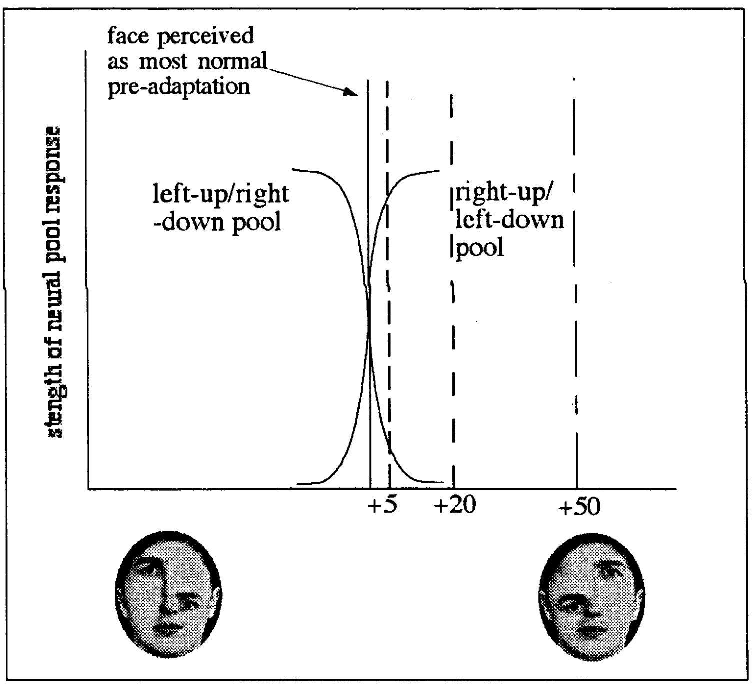


Figure 6.15. Example two-pool neural models for the (a) symmetric distortion and (b) asymmetric distortion. The relative position of the adaptors used to the coding range are shown (only positive adaptors are shown, the position of negative adaptors would be exactly reversed).

For the asymmetric distortion, it is possible to explain some of the results (but not all) in terms of a model that simply covers a smaller range of stimulus values than for symmetric. This "small range two-pool model" is illustrated in Figure 6.15b.

Assuming that the range of asymmetric eye height values coded is approximately -10 to +10 pixels-per-eye (again, see Experiment 11 for justification), the +5 pixel adaptor will be within the coding range of the two pools, while both the +20 and +50 adaptors will fall outside it. This model correctly predicts the finding of clear adaptation for asymmetric distortions with a ±5 adaptor. It also correctly predicts that the shift in the norm will be smaller for ±20 and ±50 adaptors than for ±5. However, it is only able to predict this because the 20 or 50 adaptors activate neither pool, and, under this circumstance, the model would produce no shift in the norm at all (i.e., no adaptation). This prediction is inconsistent with the finding of small but significant aftereffects for asymmetric distortions with ±20 and ±50 adaptors.

Is it possible to modify the "small range" model, by assuming that the tuning curves for each pool turn around at larger stimulus values? In Figure 6.16a, I have assumed that maximum firing in either pool will be reached quite close to the norm, but that, rather than the response dropping to zero outside this range, the curves falls off sharply at moderate deviation levels but have very extended tails. In the unadapted state, this would allow detailed perception of small differences in eye asymmetry for small deviation levels, and some degree of sensitivity to extreme asymmetric distortions, properties that would seem not unreasonable if neural resources are allocated efficiently. In terms of adaptation, the "turnaround" model is able to explain the small but nonzero adaptation for ±50 adaptors, because pre-adaptation output from the two pools is slightly different. Also, if the tuning curves had the shape indicated in the figure, it can successfully predict that adaptation for ±20 should be greater than for ±50 (because the pre-adaptation difference in pool output is larger) and less than for ±5 (because the pre-adaptation difference is smaller).

Although the "turnaround" model appears to explain the results very nicely, it has one very unfortunate drawback. Regardless of whether a difference readout or a ratio readout from the two pools is assumed, this model falsely predicts the existence of metamers (Figures 6.16b and c). The term metamer is derived from colour vision, where it refers to different physical stimuli which produce the same percept (for example, the percept of white can be produced by multiple combinations of three wavelengths). Any two-pool model that allows response curves to decrease, rather than consistently increasing towards an asymptote, will predict metamers. For example, if we assumed that response curves took the specific shape drawn, then the difference readout (Figure 6.16b) indicates that the -45 stimulus and the -5 stimulus should be perceived as identical. These face metamers clearly do not exist; indeed, two very different eye-height deviations will never look the same to observers.

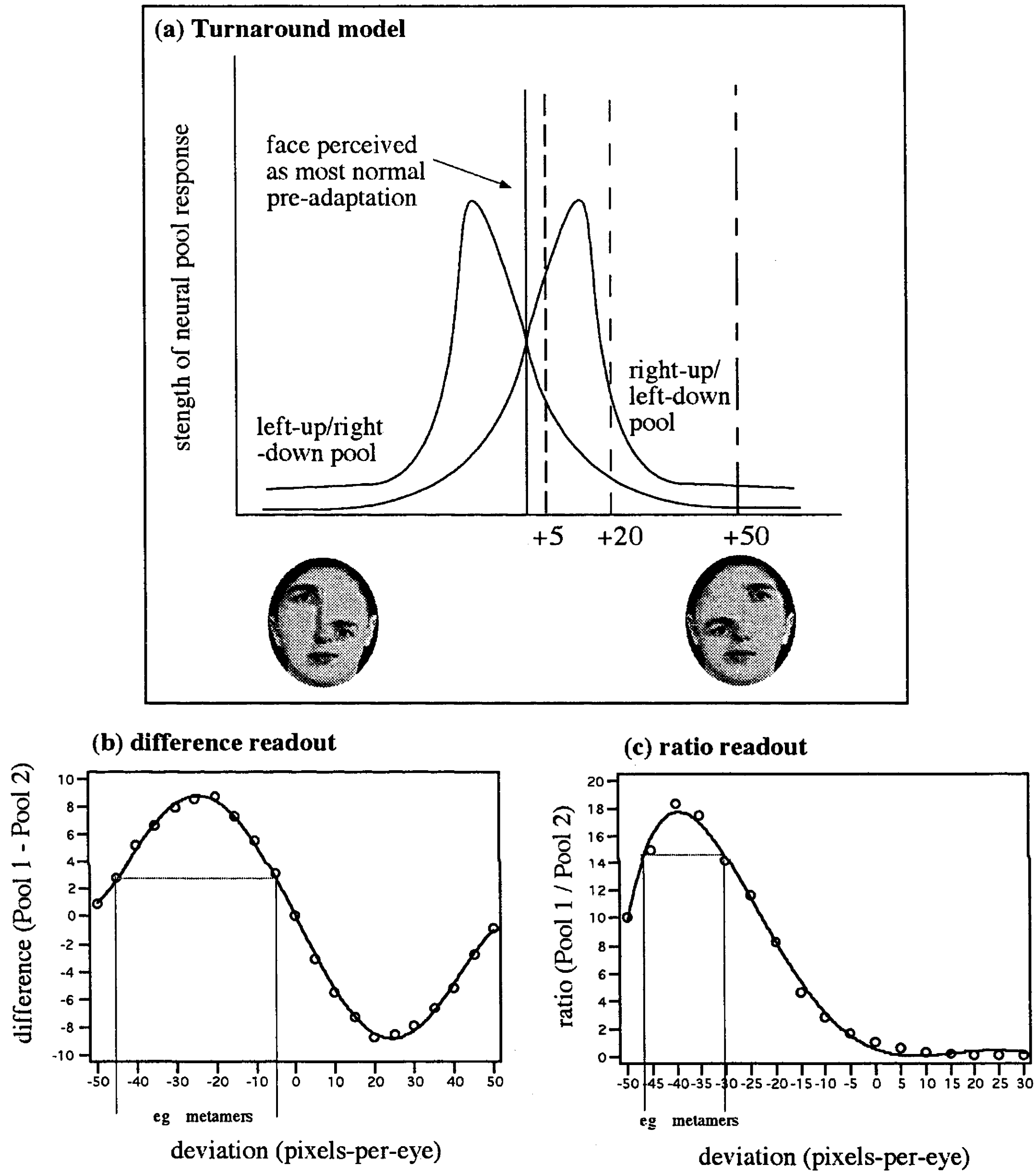


Figure 6.16. (a) The "turnaround" model shown with the positive adaptors. Also shown are (b) the difference readout and (c) the ratio readout of the model showing that in both cases metamers exist (i.e. two deviations levels which produce the same output values).

The resolution of this problem remains unclear. The standard alternative model from early vision is a series of narrow-band mechanisms. These have been used to explain effects that do not occur with respect to a single perceptual norm, such as the tilt aftereffect for line orientation, and the change in the contrast sensitivity function for spatial frequency (e.g., Blakemore & Campbell, 1969). This type of model would appear to be unsuitable here. It would predict "notch" adaptation, in which the largest change in perception occurs at the adaptor position and the pre- and post-adaptation rating curves rejoin fairly close to the adaptor location. In contrast, the current results (see Figure 6.9) showed that the greatest change in ratings always occurred partway between the original norm and the adaptor, with perception affected for a large range of deviation levels.

Given these issues, how is the problematic nonzero adaptation for moderate and extreme asymmetric adaptors to be interpreted? If it is certain that this adaptation arose from within face-space, my results argue that adaptation for asymmetric eye-height deviations cannot be explained within either of the types of standard models derived from low-/mid-level vision. This suggests that new models may need to be developed to fully explain high-level adaptation effects. However, an alternate idea is that, despite the orientation-transfer results of Experiment 12, a very small amount of adaptation for upright faces (perhaps covering the 0.34 pixel effect for ±50 adaptors) could be coming from mid-level vision mechanisms. For example, if there were a generic "equal height" or "symmetry" mechanism that could be adapted (e.g., adapting to two items with the left one lower than the right could make equal-height items appear to have the right one shifted up), then a small aftereffect for asymmetric eye height outside its coding range in face-space could potentially be explained.

Overall, I have presented a two-pool neural model that can neatly explain the coding of symmetric deviations of eye height within face-space, and the resulting patterns of adaptation. By simply reducing the range of stimulus values coded, the same model is also capable of explaining the most important difference in pattern observed for the asymmetric distortions, namely the opposite direction of adaptor position effects for asymmetric than for symmetric (although the origin of the nonzero effect for asymmetric ±20 and ±50 adaptors remains uncertain). The general success of the neural modelling approach used here argues that, in the long-term, it will be feasible to discover how specific dimensions of face-space are coded at the neural level, using adaptation results.

6.6.4 A limitation to using adaptation to explore face-space dimensions.

It is important to note that there is one question about face-space that adaptation alone cannot answer. So far, I have referred to different distortion types as corresponding to different "dimensions" of face-space, and at first glance, the distortion types chosen (e.g., height of both eyes) might seem natural candidates for "core dimensions" or "cardinal axes" of this space. However, I am not necessarily claiming that this is the case. To explain this, consider colour-space as an example. In the standard theory of colour-space, each hue is represented in terms of its position on two cardinal axes, a red-green axis and a blue-yellow axis, with white forming the perceptual norm at the centre. However, our knowledge that red-green and blue-yellow are cardinal, as opposed to purple-lime and turquoise-orange, comes from electrophysiology (e.g., De Valois, Smith, Kaitai, & Karoly, 1958), and not from adaptation studies. Although adaptation to red causes white to appear green, it is just as much the case that adaptation to purple causes white to appear lime. Thus, the mere existence of a norm-based aftereffect does not indicate that the dimension tested is cardinal, rather than being formed from a composite of the true underlying axes.

6.6.5 What do the present results tell us about the processing of inverted faces?

The theoretical discussion above applies to upright faces. While upright faces are the primary focus of this chapter, I will now briefly consider the explanation of adaptation for inverted faces (with inverted adaptors).

In the present study, results for upright and inverted faces were different in that, in the pre-adaptation ratings, the below-threshold region was wider for inverted than for upright faces (replicating McKone et al., submitted). This confirms many studies showing that sensitivity to second-order relational changes is relatively poor in inverted faces. In terms of adaptation, however, empirical results for inverted faces with inverted adaptors were similar in many ways to those for upright faces with upright adaptors (see Figure 6.12).

Clearly, adaptation for inverted faces, like that for upright faces, occurs with respect to a norm. This is true in the present research (e.g., an inverted eyes-up face made an inverted undistorted face appear eyes-down) and also in previous studies that have shown adaptation of identity in face-space (adapting to inverted anti-Tim made the inverted average face appear like Tim; Leopold et al., 2001) and adaptation of perceived sex (adapting to an inverted male face made an inverted sex-neutral face appear female; Rhodes, Jeffery et al., 2004). However, it is also apparent that the norm accessed by inverted faces must be different from that accessed by upright faces. The lack of transfer across orientations (Experiment 12) argues that the source of the aftereffects for inverted faces was different from that for upright. Also note that in Rhodes, Jeffery et al. (2004), it was shown that opposite aftereffects can simultaneously be induced to upright and inverted faces, again indicating different norms.

I suggest that inverted face adaptation might be explained as follows: inverted faces are rotated (or otherwise adjusted) to upright and compared to a norm, but within an object recognition system rather than the face system. This idea appears quite plausible, from a number of perspectives. First, the idea that inverted faces are processed within generic object system/s is a common one in the face literature, and is consistent with many empirical findings (e.g., Farah, Wilson, Drain, & Tanaka, 1995; Moscovitch, Winocur, & Behrmann, 1997; Tanaka & Farah, 1993). Second, the idea that the object system can perform rotation (e.g., Jolicoeur, 1985; Lawson & Jolicoeur, 1998; McKone & Grenfell, 1999), but the face system cannot, is also consistent with evidence that the holistic component of face processing cannot be learned for inverted faces even with extensive practice (Chapter 3; McKone, Martini & Nakayama, 2001). Third, the claim that the object system can form norms is perfectly reasonable, given that the ability to form prototypes is a standard part of all human categorisation processes (e.g., Williams, Fryer, & Aiken, 1977). Finally, the current finding that, in the unadapted state, sensitivity to eye height deviations was weaker for inverted faces than for upright faces is consistent with a common idea that, while the face system codes very detailed information about distances between parts, object processing codes only coarse relations (e.g., Beiderman, 1987).

6.6.6 Conclusion.

The present study has demonstrated that some dimensions of face-space are more adaptable than others and that these effects for upright faces can be attributed to processing specifically within the face recognition system. I have also provided a neural model of the adaptation effects, including those of adaptor position, that works well for symmetric distortions of eye height, if not entirely so for asymmetric distortions. Taken together, the empirical results and theoretical ideas I have introduced suggest that future adaptation studies will provide much valuable insight into the structure of face-space. A particularly valuable approach is to compare adaptation for different types of simple distortions, corresponding to displacement along different trajectories in face-space.

In terms of future research, testing both additional relational distortions (e.g., nose shifts, mouth shifts) and also local feature distortions (e.g., wide vs. narrow eyes, or thin vs. fat lips), are likely to be useful. Testing adaptation to relational-only or featural-only changes, in both upright and inverted orientations, would be of particular interest, given the literature suggesting that perception of relational changes is more affected by inversion than perception of local feature changes. Theoretically, studies of this type may begin to bring together the two major, but currently largely independent, theoretical approaches in face recognition. The concept of face-space has been used to explain the representation of individual identity, and also distinctiveness, attractiveness, caricature and cross-race effects (see Chapter 2). Independently, the concept of holistic/configural/relational processing has been used to explain differences between upright and inverted faces, and between faces and objects. To date, only a few studies (e.g., Leder & Bruce, 1998) have made reference to both ideas simultaneously, or attempted to theoretically integrate these two important ideas.

CHAPTER 7: FACES, OBJECTS AND EXPERIENCE: THEORETICAL CONCLUSIONS AND EMPIRICAL ADVANCES

This chapter forms the General Discussion for the thesis as a whole. The empirical work (Chapters 3-6) has been presented as four more-or-less independent papers, with substantial theoretical discussion of the results included within each chapter. The details of those discussions will not be repeated here. Instead, the role of this chapter is to briefly draw together the major findings of the thesis, and put them back into the context of the general theoretical questions raised in the Introduction (Chapters 1 and 2).

Chapter 7 begins with a summary of the major empirical results and methodological advances made in the thesis (Section 7.1). Turning to theoretical questions, in Section 7.2, I then consider the results in relations to the two aims of the thesis stated at the beginning of Chapter 1, namely to contribute to the debate on whether holistic/configural processing is specific to the domain of faces, and to use the effects of experience to explore the representations of faces as individuals within face-space. Following this, in Section 7.3, I link the present results with more general previous literature to address the question raised in the title of the thesis, namely "What changes with experience for face and object processing?".

7.1 New empirical findings from this thesis

The major new empirical results of the thesis fall into four groups. These correspond to the four experimental chapters.

In Chapter 3, I showed that holistic/configural processing does <u>not</u> develop for inverted faces with practice. Despite using stimuli designed to encourage maximum reliance on holistic information – multiple photographs of identical twins in different views – extensive experimental training (8 hrs) with inverted faces produced no evidence of holistic processing. There was no composite effect for inverted faces, and subjects who had learned to identify the twins inverted did so almost entirely by developing very local feature strategies (e.g., a difference in eyebrow grooming) that would not provide a viable means of achieving face recognition in real world settings.

In Chapter 4, in preparation for testing dog experts, I examined the performance of dog novices on three tasks, each of which contrasted results for face stimuli with results for labrador stimuli. Despite the requirement for within-class discrimination in every task, and the fact that dogs provide a stimulus class well-matched to faces (e.g., they have genetic variability, fuzzy part boundaries, etc), no evidence of face-like processing was found. Results of a recognition memory task replicated the standard finding of a much larger inversion effect for faces than for dogs. A contrast reversal test replicated previously found substantial effects of contrast reversal for faces, but found only a very weak effect for dogs; this experiment provided the first test of contrast reversal effects for a natural object class for which shape-from-shading information is potentially useful. The composite task replicated the standard composite effect for faces, but found no effect for dogs; this provided the first test of the composite task for any class of natural object. In addition to the new empirical results, this chapter produced an important methodological contribution, namely the evidence that the composite task provides a purer measure of the type of holistic/configural processing used for faces than other standard tasks: these tasks (inversion, part-whole) produce effects for objects that are smaller for objects than for faces but are still greater than zero, while the composite task has now been shown to produce <u>no</u> effect for both a natural object class (dogs), and, previously, for an artificial object class (greebles).

In Chapter 5, dog experts were tested on the tasks of Chapter 4. Experts had 5-37 yrs experience with labradors (mean = 23 yrs), and there was evidence of genuine expertise both anecdotally (e.g., experts commented on which dogs were US-bred and which Australian-bred), and behaviourally (e.g., the older experts showed poorer performance on faces than the younger novices, but better performance on dogs). Despite this, findings for experts looking at dog stimuli exactly matched those of novices – that is, a small inversion effect on memory, a weak contrast reversal effect, and no composite effect – rather than the patterns observed in both subject groups with face stimuli. Thus, even with substantial real-world expertise, processing of objects remained part-based rather than holistic.

In Chapter 6, I examined adaptation aftereffects for upright faces within face-space. Results showed that it was easier to temporarily shift the norm of face-space in some directions than in others, with more adaptability along a dimension associated with large variability, and less adaptability along a dimension associated with small variability. For adaptors placed a long way from the original norm, symmetric distortions of eye height showed a larger aftereffect than asymmetric distortions of eye height, and symmetric distortions also produced strong adaptation for a wider range of adaptor positions than did asymmetric distortions. This study is the first to use simple relational distortions, and to contrast adaptability for different relational changes. Results suggest that this approach, as opposed to using more global distortions, provides a valuable new method for investigating the structure of face-space.

7.2 Revisiting the aims of the thesis

The theoretical questions addressed by each empirical chapter can be grouped into two overall aims. The first of these was to investigate the theory of domain-specificity for faces; the second was to investigate the representation of individual faces within face-space.

7.2.1 Domain-specificity for faces, versus alternative ideas.

The domain-specificity view states that faces *per se* are "special" in comparison to other objects, and moreover that it is particularly upright faces that are special. This is claimed to involve specialised neural processing for faces (i.e., involving the Fusiform Face Area), and also a specialised cognitive processing style (i.e., holistic/configural processing). Two major alternatives to the domain-specificity idea have been put forward in the literature: the within-class discrimination hypothesis, and the expertise hypothesis. Both can be assessed by comparing upright faces to inverted faces and by comparing faces to non-face objects.

Many results in the present thesis are consistent with the standard position that upright faces are special compared to inverted faces. The inversion effect on recognition memory was much larger for faces than for dogs (Chapter 4), replicating many studies beginning with Yin (1969). The composite effect, indicating holistic processing, was present for upright faces but not inverted faces (Chapters 3, 4, 5), replicating Young, Hellawell and Hay (1987) and others. Sensitivity to second-order relational changes was greater for upright faces than for inverted faces (Chapter 6), also replicating many previous findings (e.g., Leder & Bruce, 1998); although no direct comparison to a local feature change was included, this result is at least consistent with the idea that spacing information forms a key part of the holistic/configural processing for upright faces (e.g., Maurer, Le Grand and Mondloch, 2002). In addition to these results indicating special processing at a behavioural level, evidence of special processing at the neural level was also obtained. In the adaptation study (Chapter 6), there was no transfer of adaptation across orientations, arguing that, consistent with the findings of Rhodes, Jeffery et al. (2004), upright and inverted faces are processed by independent groups of cells.

The next question was whether faces are special in comparison to <u>other objects</u>, when both stimulus classes are shown upright. Results here clearly supported domain-specificity.

The idea that task requirements for <u>within-class discrimination</u> (particularly, discrimination at the individual level) would be sufficient to produce "special" processing for objects had already been largely disproved by previous literature (see all of Chapter 1). The results of Chapter 4 confirmed this conclusion: when subjects had no particular expertise with the stimulus class, processing style for dogs did not match that for faces in three tests requiring reference to individual identity (inversion effects on memory, contrast reversal, and composite task).

Most authors accept that the within-class discrimination hypothesis is no longer viable (although see Tarr, 2003), but many see the <u>expertise hypothesis</u> as a likely alternative to domain-specificity, or even as having been clearly demonstrated to be correct (see Discussion of Chapter 5). According to the expertise hypothesis, faces appear "special" only due to normal adults' expertise at within-class discrimination for faces compared to their lack of expertise at within-class discrimination for other objects. The expertise hypothesis predicts that any class of object should also show the hallmarks of face-like processing, when all exemplars share a first-order configuration and when the subject has sufficient expertise to take advantage of minor second-order differences between exemplars (Diamond & Carey, 1986).

The literature review in Chapter 5 made it clear that the evidence previously interpreted as favouring the expertise hypothesis is in fact much weaker than commonly assumed, and that many studies do not support it at all. My own results also argue against it. The lack of learning of holistic processing for inverted faces (Chapter 3) is of some relevance, in that the amount of training used (8 hrs) was as much as in the studies claiming to have found holistic processing for greebles (e.g., Gauthier & Tarr, 1997, 2002; Gauthier, Williams, Tarr & Tanaka, 1998). Of course, this study alone does not disprove the expertise hypothesis: greater experience might be required, or the previous existence of holistic processing in the upright orientation might interfere with its acquisition for inverted faces. The lack of holistic processing for dogs in dog experts, however, is compelling evidence against the expertise hypothesis. In this case, the experts in question had many years' experience at making individual-level discriminations of the particular breed of dogs used as stimuli (labradors), and substantially more experience than that required to produce holistic processing of faces in children (at most 4 years, and possibly much less; see Chapter 2). Despite this, no holistic processing emerged, even in the upright orientation, namely the orientation with which the experts' lifetime experience had been gained.

Overall, this thesis has argued that neither the within-class discrimination hypothesis nor the expertise hypothesis can be supported. Instead, it seems that it is faces *per se* that are "special".

7.2.2 Possible origins of domain-specificity for faces.

While my empirical results support domain-specificity for faces, my experiments have not addressed the origin of this domain-specificity. Given the evidence that expertise obtained in late childhood or adulthood with objects does not produce the same type of processing as occurs naturally for faces, it would seem that there is either an innate component to holistic processing for faces, and/or that something important occurs in infancy or early childhood to establish it.

A review of the possible origins was provided in Chapter 2. Briefly, there is some evidence that newborn babies have a preference for upright-face-like stimuli (e.g., Johnson, Dziurawiec, Ellis, & Morton, 1991), although not everyone is convinced that it is the face-like aspects of the stimuli to which infants are responding in these tests (e.g., Cassia, Turati, & Simion, 2004). There is also some evidence suggesting that, while 7 month infants can discriminate individual faces well (Fagan, 1979), at the same age they are poor at discriminating between non-face objects (Bonatti, Frot, Zangl & Mehler, 2002), although well-controlled studies in this area (e.g., individual face discrimination versus individual dog discrimination) do not appear to have been run. Studies such as these are consistent with the suggestion by Johnson and colleagues that there is an innate representation of basic face structure (they propose this is located subcortically), which biases newborns' attention towards faces, and drives rapid development of the ability to distinguish between individual people based on their faces. It is also clear that any such innate representation requires appropriate visual input during a critical period in early infancy (very approximately, before the first 3 months of life), given the studies showing that patients with congenital bilateral or left-eye cataracts fail to develop holistic face processing even after many years of exposure to faces after removal of the cataracts (Le Grand, Mondloch, Maurer, & Brent, 2001, 2003, 2004).

While the idea of some innate component to "special" face recognition is a common one, I have also noted (Chapter 2, Chapter 5), that it is not the only possibility. It may instead be that the cortical area associated with face recognition in adults has, at the beginning of life, the capacity to learn holistic processing for any stimulus class of which enough exemplars are seen during the critical period. Given that most parents would be unwilling to bring up their newborn infant with close-up exposure to as many different dogs (for example) as the infant sees faces, this hypothesis cannot be ethically tested in humans. In terms of future research, it would be interesting to know whether baby monkeys raised primarily looking at a class of non-face objects, or at inverted faces (in an otherwise upright world), would develop evidence of face-like processing for these stimuli. (Note that this test would require first demonstrating that monkeys show the same hallmarks of holistic processing as do humans for upright faces, such as the composite effect.)

Currently, the origins of "special" processing for faces remain unclear. The fact that two reasonable mechanisms for domain-specificity in adults can be outlined, however, demonstrates the theoretical viability of the view. Also note that, even if development during early infancy of face-like processing for other objects were discovered to be possible under specialised circumstances (e.g., monkeys raised with pictures of dogs rather than with monkey/human faces), these circumstances would be extremely unlikely to ever occur for a human baby, given the small chance of seeing appropriate stimuli during the critical period. Thus, to all intents and purposes, face-recognition in humans would remain domain-specific.

7.2.3 Understanding the representations of individual (upright) faces.

The second aim of this thesis was contribute to our understanding of how faces themselves are represented, focussing on norm-based coding and the theory of face-space. Addressing this aim formed a smaller part of the thesis than addressing domain-specificity. However, some interesting conclusions may be drawn here too. My results (Chapter 6) supported previous ideas that faces are coded with respect to a norm, or average face, and that adaptation can shift perception relative to this norm. Moreover, I have shown that, while the norm of face-space is quite changeable in the short term, it is changeable only in specific ways that are affected by long-term experience. In particular, the amount by which the norm can be shifted (measured both by the range of adaptors which produce an aftereffect and the size of the aftereffect at specific adaptor positions) is related to the range of values (e.g., eye-heights) previously experienced on a particular dimension. This is an important and novel conclusion.

In terms of future research, it would be interesting to test adaptation for various objects. Some kind of adaptation for objects seems entirely likely. Even more interesting would be to see whether dissociations between populations of neurons coding objects and upright faces could be found, similar to the dissociation which has been shown between upright and inverted faces (Chapter 6; Rhodes, Jeffery et al., 2004).

7.3 What changes with experience for face and object processing?

At the most general level, both topics of the present thesis – domain-specificity of face recognition and face-space coding of individuals – are related via a general question about the effects of experience on face and object processing. The timescales of this experience are, of course, quite different in each case: domain-specificity is addressed by asking whether face-like "special" processing can develop for other stimuli with many hours or years of practice, while face-space coding is addressed by examining the effects of a few minutes or seconds of adaptation. To draw these two parts of the thesis together, in this final section I consider the results in terms of summarising what does, and does not, change with experience for upright faces, inverted faces, and non-face objects. For each stimulus type, I consider the effects of experience on holistic processing, on norm-based coding/face-space, and on other important aspects of processing not specifically tested in the present thesis; I also consider timescales ranging from minutes to lifetime exposure.

7.3.1 What changes with experience for upright faces?

One of the most important aspects of upright face processing – holistic processing – seems to change very little with experience, at least after early infancy.

Holistic processing is present in 4-year old children, the youngest age-group comprehensively tested (e.g., McKone & Boyer, submitted; Pellicano & Rhodes, 2003). There is also suggestive evidence of its presence in 7-month olds (integration across internal and external features, Cohen & Cashon, 2001). However, although it is clear that extensive lifetime experience is not needed to develop holistic processing for upright faces, there is evidence that experience with faces is needed during the critical period in early infancy (Le Grand et al., 2001, 2003, 2004).

Discrimination of individual faces is also good from an early age. For example, newborns are able to discriminate their mother from a stranger at 4 hours old (Pascalis, de Schonen, Morton, Deruelle, & Fabregrenet, 1995), and by 7 months old infants can discriminate two similar unknown faces (Fagan, 1979). Experience can act to de-tune the ability to discriminate individuals of another species. Infants do not show a cross-species deficit (discriminating monkey faces vs. discriminating human faces) at 6 months old, but they do at 9 months old. This is similar to the finding that newborns can discriminate all phonemes but that, with experience, they loose the ability to discriminate those phonemes not present in the native language to which they are exposed (e.g., Werker & Tees, 1984).

With respect to face-space, it is likely that further tuning of the dimensions used to individuate faces occurs throughout life. Cross-race deficits occur when the subject is unfamiliar with another race, but seem to be smaller with more experience of that race (e.g., Rhodes, 1993) even when that experience is gained as an adult (e.g., Tanaka, Kiefer, & Bukach, 2004).

With respect to short-term changes in face-space, representations of norm/s may also change in response to short-term experience (e.g., Leopold, O'Toole, Vetter, & Blanz, 2001; Rhodes, Jeffery, Watson, Clifford, & Nakayama, 2003; Webster & MacLin, 1999). This thesis has shown that they do so in ways based on previous long-term experience. In particular, the amount of adaptation possible (both in terms of the size of the aftereffect and the range of values at which adaptation occurs) is linked to the variability or range which a particular dimension needs to be able to code for real faces.

7.3.2 What changes with experience for inverted faces?

With respect to holistic processing, subjects do not holistically process inverted faces, instead using part-based processing, and this style of processing does not change after extensive experimental practice (Chapter 3; see also McKone, Martini, & Nakayama, 2001). Of course, such experience is gained as an adult; that is, after the end of the critical period for developing holistic processing for upright faces. It is possible that if an infant were raised seeing only inverted faces during the critical period there would be holistic processing for these instead of upright faces.

Norm-based coding seems to occur for inverted faces as well as for upright faces, although evidence suggests that different norms are used for each (Chapter 6; Rhodes, Jeffery et al., 2004). It seems likely that inverted faces are processed in the object system where coding of relational (spacing) information is much more coarse than the coding of relational information available for representations of upright faces. The development of a norm for inverted faces, like that for upright faces, presumably comes from averaging lifetime experience. Moreover, like the norm for upright faces, the norm for inverted faces can be temporarily shifted by adaptation (Chapter 6; Leopold et al., 2001; Rhodes, Jeffery et al., 2004; Webster & MacLin, 1999).

7.3.3 What changes with experience for objects?

There are many interesting questions pertaining to objects and experience, here I have been interested in only a sub-set of these. In particular I have looked at the effects of experience on object perception in the context of possible developments of face-like processing.

With respect to holistic processing, even after many years of experience, objects are not processed in the same way as upright faces. Instead processing style remains the same as in novices (i.e., part-based), although, as with inverted faces, people become better at using this style to differentiate individual exemplars of the stimulus class (Chapter 5). When discussing object processing in this context it is important to note the particular use of the term "holistic" in the face recognition literature. There is no doubt that objects are in some sense processed in terms of the whole, even without extensive experience. For example, small part-whole effects are found for a range of objects (Davidoff & Donnelly, 1990; Donnelly & Davidoff, 1999; Gauthier & Tarr, 1997; Tanaka et al., 1996, cited in Tanaka & Gauthier, 1997). Similarly, object superiority effects exist for discriminating lines in the context of more complex shapes (e.g., Enns & Gilani, 1988). However, I believe it is clear from the studies reviewed and conducted in this thesis that such "whole-based" processing is not the same as the holistic/configural processing used for upright faces, despite possible confusions caused by the terminology.

With respect to norm-based coding, it is clear that prototypes (i.e., norms) can form for objects (e.g., Williams, Fryer, & Aiken, 1977). Similarly, objects which are more average are seen as more attractive (Halberstadt & Rhodes, 2000, 2003). Caricature effects are also found for bird experts looking at very similar birds; that is caricatures were recognised more quickly than anti-caricatures (Rhodes & McLean, 1990). These findings argue that, as for faces, norms for objects can be developed that reflect the range of real world experience. Whether short-term adaptation effects occurs for objects norms has not been tested.

Finally, experience may, of course, change object processing in other ways. In one well-known effect, experience at recognising inverted objects seems to remove inversion effects on naming time with little practice (McKone & Grenfell, 1999; Tarr & Pinker, 1989; as noted above, inversion effects for faces remain even after extensive practice). Experience may also cause a downward shift in the level at which objects are first categorised, so that an expert bird-watcher would be more likely to say "robin" rather then "bird" when presented with a robin (K. E. Johnson & Mervis, 1997; Tanaka & Taylor, 1991). While this is similar to faces (Tanaka, 2001) the present thesis has made it clear that objects are not processed like faces in many other ways.

Anglin, A. D., Hser, Y., & Chou, C. (1993). Reliability and validity of retrospective behavioural self-report by narcotics addicts. *Evaluation Review, 17*(1), 91-108.

Ashbridge, E., Perrett, D. I., Oram, M. W., & Jellema, T. (2000). Effect of image orientation and size on object recognition: Responses of single units in the macaque monkey temporal cortex. *Cognitive Neuropsychology, 17*(1-3), 13-34.

Atick, J. J., Griffin, P. A., & Redlich, A. N. (1996). The vocabulary of shape: Principal shapes for probing perception and neural response. *Network-Computation in Neural Systems, 7*(1), 1-5.

Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive Psychology, 25*(3), 281-316.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience, 8*(6), 551-565.

Benton, A. (1990). Facial Recognition 1990. *Cortex, 26*(4), 491-499.

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review, 94*(2), 115-147.

Blakemore, C., & Campbell, F. W. (1969). Adaptation to Spatial Stimuli. *Journal of Physiology, 200*(1), P11-P13.

Blakemore, C., & Cooper, G. F. (1970). Development of the brain depends on the visual environment. *Nature, 228*(31), 477-478.

Bonatti, L., Frot, E., Zangl, R., & Mehler, J. (2002). The human first hypothesis: identification of conspecifics and individuation of objects in the young infant. *Cognitive Psychology, 44*(4), 388-426.

Bradshaw, J. L., & Wallace, G. (1971). Models for the processing and identification of faces. *Perception and Psychophysics, 9*(5), 443-448.

Broad, K. D., Mimmack, M. L., & Kendrick, K. M. (2000). Is right hemisphere specialization for face discrimination specific to humans? *European Journal of Neuroscience, 12*(2), 731-741.

Brown, V., Huey, D., & Findlay, J. M. (1997). Face detection in peripheral vision: do faces pop out? *Perception, 26*(12), 1555-1570.

Bruce, V., Burton, M., & Craw, I. (1992). Modelling face recognition. *Philosophical Transcripts of the Royal Society London: B, 335*, 121-128.

Bruce, V., Doyle, T., Dench, N., & Burton, M. (1991). Remembering Facial Configurations. *Cognition, 38*(2), 109-144.

Bruce, V., & Langton, S. (1994). The Use of Pigmentation and Shading Information in Recognizing the Sex and Identities of Faces. *Perception, 23*(7), 803-822.

Bruce, V., & Young, A. (1986). Understanding Face Recognition. *British Journal of Psychology, 77*, 305-327.

Bruyer, R., & Crispeels, G. (1992). Expertise in Person Recognition. *Bulletin of the Psychonomic Society, 30*(6), 501-504.

Burgund, E. D., & Marsolek, C. J. (2000). Viewpoint-invariant and viewpoint-dependent object recognition in dissociable neural subsystems. *Psychonomic Bulletin and Review, 7*(3), 480-489.

Burton, A. M., & Vokey, J. R. (1998). The face-space typicality paradox: Understanding the face-space metaphor. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 51*(3), 475-483.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science, 195*(4275), 312-314.

Carey, S., & Diamond, R. (1994). Are faces perceived as configurations more by adults than by children? *Visual Cognition, 1*(2/3), 253-274.

Cassia, V. M., Turati, C., & Simion, F. (2004). Can a nonspecific bias toward top-heavy patterns explain newborns' face preference? *Psychological Science, 15*(6), 379-383.

Clifford, C., Wenderoth, P., & Spehar, B. (2000). A functional angle on some after-effects in cortical vision. *Procedings of the Royal Society London, B, 267*, 1705-1710.

Cohen, L. B., & Cashon, C. H. (2001). Do 7-month-old infants process independent features or facial configurations? *Infant & Child Development, 10*(1-2), 83-92.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments and Computers, 25*, 257-271.

Collishaw, S. M., & Hole, G. J. (2002). Is there a linear or a nonlinear relationship between rotation and configural processing of faces? *Perception, 31*(3), 287-296.

Cooper, E. E., & Brooks, B. E. (2004). Qualitative differences in the representation of spatial relations for different object classes. *Journal of Experimental Psychology: Human Perception and Performance, 30*(2), 243-256.

Cooper E. E., & Wojan, T. J (2000). Differences in the coding of spatial relations in face identification and basic-level object recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*(2), 470-488.

Damasio, A. R., Damasio, H., & Van Hoesen, G. W. (1982). Prosopagnosia: anatomic basis and behavioral mechanisms. *Neurology, 32*(4), 331-341.

Davidoff, J., & Donnelly, N. (1990). Object superiority: A comparison of complete and part probes. *Acta Psychologica, 73*(3), 225-243.

de Gelder, B., Bachoud-Levi, A. C., & Degos, J. D. (1998). Inversion superiority in visual agnosia may be common to a variety of orientation polarised objects besides faces. *Vision Research, 38*(18), 2855-2861.

de Gelder, B., & Rouw, R. (2000). Paradoxical configuration effects for faces and objects in prosopagnosia. *Neuropsychologia, 38*(9), 1271-1279.

de Gelder, B., & Rouw, R. (2001). Beyond localisation: a dynamical dual route account of face recognition. *Acta Psychologia, 107*(1-3), 183-207.

de Haan, M., Humphreys, K., & Johnson, M. H. (2002). Developing a brain specialized for face perception: A converging methods approach. *Developmental Psychobiology, 40*(3), 200-212.

De Renzi, E. (1986). Current issues on prosopagnosia. In H. Ellis, M. A. Jeeves, F. Newcombe & A. Young (Eds.), *Aspects of face processing* (Vol. 28, pp. 243-252.). Dordrecht: Martinus Nijhoff Publishers.

De Valois, R. L., Smith, C. J., Kaitai, S. T., & Karoly, A. J. (1958). Response of single cells of monkey lateral geniculate nucleus to monochromatic light. *Science, 127*, 238-239.

Deng, C., & Rogers, L. J. (2002). Social recognition and approach in the chick: lateralization and effect of visual experience. *Animal Behaviour, 63*, 697-706.

Diamond, R., & Carey, S. (1977). Developmental changes in the representation of faces. *Journal of Experimental Child Psychology, 23*(1), 1-22.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of expertise. *Journal of Experimental Psychology: General, 115*(2), 107-117.

Dixon, M. J., Desmarais, G., Gojmerac, C., Schweizer, T. A., & Bub, D. N. (2002). The role of premorbid expertise on object identification in a patient with category-specific visual agnosia. *Cognitive Neuropsychology, 19*(5), 401-419.

Donnelly, N., & Davidoff, J. (1999). The mental representations of faces and houses: Issues concerning parts and wholes. *Visual Cognition, 6*(3-4), 319-343.

Donovan, J. D., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology, 84*(5), 795-805.

Downing, P. E., Jiang, Y. H., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science, 293*(5539), 2470-2473.

Duchaine, B. C. (2000). Developmental prosopagnosia with normal configural processing. *Neuroreport, 11*(1), 79-83.

Duchaine, B. C., Dingle, K., Butterworth, E., & Nakayama, K. (2004). Normal greeble learning in a severe case of developmental prosopagnosia. *Neuron, 43*(4), 469-473.

Duchaine, B. C., Nieminen-von Wendt, T., New, J., & Kulomaki, T. (2003). Dissociations of visual recognition in a developmental agnosic: evidence for separate developmental processes. *Neurocase, 9*(5), 380-389.

Elgar, K., & Campbell, R. (2001). The development of face-identification skills: What lies behind the face module? *Infant & Child Development, 10*(1-2), 25-30.

Endo, M., Masame, K., & Maruyama, K. (1990). A limited use of configural information in the perception of inverted faces. *Tohoku Psychologica Folia, 49*, 114-125.

Enns, J. T., & Gilani, A. B. (1988) Three-dimensionality and discriminability in the object-superiority effect. *Perception and Psychophysics, 44*, 243-256.

Fagan, J. F. (1979). The origins of facial pattern recognition. In M. Bornstein & W. Keesen (Eds.), *Psychological development from infancy: Image to intention* (pp. 83-113.). Hillsdale, New Jersey: Erlbaum.

Fallshore, M., & Schooler, J. W. (1995). Verbal vulnerability of perceptual expertise. *Journal of Experimental Psychology: Learning Memory and Cognition, 21*(6), 1608-1623.

Farah, M. J. (1996). Is face recognition 'special'? Evidence from neuropsychology. *Behavioural Brain Research, 76*(1-2), 181-189.

Farah, M. J., Levinson, K. L., & Klein, K. L. (1995). Face perception and within-category discrimination in prosopagnosia. *Neuropsychologia, 33*(6), 661-674.

Farah, M. J., Rabinowitz, C., Quinn, G. E., & Liu, G. T. (2000). Early commitment of neural substrates for face recognition. *Cognitive Neuropsychology, 17*(1-3), 117-123.

Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. W. (1995). The inverted face inversion effect in prosopagnosia: evidence for mandatory, face-specific perceptual mechanisms. *Vision Research, 35*(14), 2089-2093.

Farkas, L. G., Hreczko, T. A., & Katic, M. J. (1994). Craniofacial norms in North American Caucasians from birth (one year) to young adulthood. In L. G. Farkas (Ed.), *Anthropometry of the head and face* (2nd ed., pp. 241-335). New York: Raven Press.

Fastenau, P. S., Denburg, N. L., & Abeles, N. (1996). Age Differences in Retrieval: Further Support for the Resource-Reduction Hypothesis. *Psychology and Aging, 11*(1), 140-146.

Fechner, G. (1966:1860). *Elements of Psychophysics, Vol. 1.* New your: Holt, Rinehart and Winston.

Galper, R. E. (1970). Recognition of faces in photographic negative. *Psychonomic Science, 19*(4), 207-208.

Gauthier, I., Anderson, A. W., Tarr, M. J., Skudlarski, P., & Gore, J. C. (1997). Levels of categorization in visual recognition studied using functional magnetic resonance imaging. *Current Biology, 7*(9), 645-651.

Gauthier, I., Behrmann, M., & Tarr, M. J. (1999). Can face recognition really be dissociated from object recognition? *Journal of Cognitive Neuroscience, 11*(4), 349-370.

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience, 6*(4), 428-432.

Gauthier, I., & Nelson, C. A. (2001). The development of face expertise. *Current Opinion in Neurobiology, 11*, 219-224.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*(2), 191-197.

Gauthier, I., & Tarr, M. J. (1997). Becoming a "greeble" expert: Exploring mechanisms for face recognition. *Vision Research, 37*(12), 1673-1682.

Gauthier, I., & Tarr, M. J. (2002). Unravelling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception & Performance, 28*(2), 431-446.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience, 2*(6), 568-573.

Gauthier, I., Tarr, M. J., Moylan, J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (2000). Does visual subordinate-level categorisation engage the functionally defined fusiform face area? *Cognitive Neuropsychology, 17*(1-3), 143-163.

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training 'greeble' experts: a framework for studying expert object recognition processes. *Vision Research, 38*(15-16), 2401-2428.

Geldart, S., Maurer, D., & Henderson, H. (1999). Effects of the height of the internal features of faces on adults' aesthetic ratings and 5-month-olds' looking times. *Perception, 28*(7), 839-850.

Geldart, S., Mondloch, C. J., Maurer, D., de Schonen, S., & Brent, H. P. (2002). The effect of early visual deprivation on the development of face processing. *Developmental Science, 5*(4), 490-501.

Gilchrist, A., & McKone, E. (2003). Early maturity of face processing in children: Local and relational distinctiveness effects in 7-year-olds. *Visual Cognition, 10*(7), 769-793.

Golby, A. J., Gabrieli, J. D., Chiao, J. Y., & Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other- race faces. *Nature Neuroscience, 4*(8), 845-850.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience, 7*(5), 555-562.

Haggbloom, S. J., & Warnick, J. E. (2003). Practice and incentive motivation in recognition of inverted faces. *Perceptual and motor skills, 96*, 578-588.

Halberstadt, J., & Rhodes, G. (2000). The attractiveness of nonface averages: implications for an evolutionary explanation of the attractiveness of average faces. *Psychological Science, 11*(4), 285-289.

Halberstadt, J., & Rhodes, G. (2003). It's not just average faces that are attractive: Computer- manipulated averageness makes birds, fish, and automobiles attractive. *Psychonomic Bulletin & Review, 10*(1), 149-156.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*(6), 223-233.

Hellige, J. B., & Michimata, C. (1989). Categorization versus distance: hemispheric differences for processing spatial information. *Memory and Cognition, 17*(6), 770-776.

Hirsch, H., & Spinelli, D. N. (1970). Visual experience modifies distribution of horizontally and vertically oriented receptive fields in cats. *Science, 168*, 869-871.

Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. *Perception, 23*(1), 65-74.

Hole, G. J., George, P. A., & Dunsmore, V. (1999). Evidence for holistic processing of faces viewed as photographic negatives. *Perception, 28*(3), 341-359.

Hreczko, T. A., & Farkas, L. G. (1994). Norms of the craniofacial asymmetries in North American Caucasians. In L. G. Farkas (Ed.), *Anthropometry of the head and face* (2nd ed., pp. 359-380.). New York: Raven Press.

Hsu, S.-M., & Young, A. W. (2004). Adaptation effects in facial expression recognition. *Visual Cognition, 11*(7), 871-899.

Humphreys, G. W., & Rumiati, R. I. (1998). Agnosia without prosopagnosia or alexia: Evidence for stored visual memories specific to objects. *Cognitive Neuropsychology, 15*(3), 243-277.

Husk, J. S., Sekuler, A. B., & Bennett, P. (2004). Specificity of inversion effects in perceptual learning. Paper presented at the Visual Sciences Society Annual Meeting, Sarasota, Florida, USA.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40*(1-2), 1-19.

Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General, 126*(3), 248-277.

Johnston, A., Hill, H., & Carman, N. (1992). Recognizing Faces - Effects of Lighting Direction, Inversion, and Brightness Reversal. *Perception, 21*(3), 365-375.

Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory and Cognition, 13*(4), 289-303.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience, 3*(8), 759-763.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*(11), 4302-4311.

Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition, 68*(1), B1-11.

Keane, S. K., Hayward, W. G., & Burke, D. (2003). Detection of three types of changes to novel objects. *Visual Cognition, 10*(1), 101-127.

Kemp, R., McManus, C., & Pigott, T. (1990). Sensitivity to the Displacement of Facial Features in Negative and Inverted Images. *Perception, 19*(4), 531-543.

Kemp, R., Pike, G., White, P., & Musselman, A. (1996). Perception and recognition of normal and negative faces: The role of shape from shading and pigmentation cues. *Perception, 25*(1), 37-52.

Kendrick, K. M., Atkins, K., Hinton, M. R., Heavens, P., & Kevern, B. (1996). Are faces special for sheep? Evidence from facial and object discrimiantion learning tests showing effects of inversion and social familiarity. *Behavioural Processes, 38*, 19-35.

Kosslyn, S. M., Koenig, O., Barrett, A., Cave, C. B., Tang, J., & Gabrieli, J. D. (1989). Evidence for two types of spatial representations: hemispheric specialization for categorical and coordinate relations. *Journal of Experimental Psychology: Human Perception and Performance, 15*(4), 723-735.

Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science, 293*(5534), 1506-1509.

Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: evidence from a parametric fMRI study. *NeuroImage, 21*(3), 946-953.

Lawson, R., & Jolicoeur, P. (1998). The effects of plane rotation on the recognition of brief masked pictures of familiar objects. *Memory and Cognition, 26*(4), 791-803.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Early visual experience and face processing. *Nature, 410*, 890.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2003). Expert face processing requires visual input to the right hemisphere during infancy. *Nature Neuroscience, 6*(10), 1108-1112.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2004). Impairment in holistic face processing following early visual deprivation. *Psychological Science, 15*(11), 762-768.

Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *Quarterly Journal of Experimental Psychology: A, 51*(3), 449-473.

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level after effects. *Nature Neuroscience, 4*(1), 89-94.

Leopold, D. A., Rhodes, G., Muller, K., & Jeffery, L. (in press). The dynamics of visual adaptation to faces.

Lewis, M. B., & Johnston, R. A. (1997). The Thatcher illusion as a test of configural disruption. *Perception, 26*(2), 225-227.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: an MEG study. *Nature Neuroscience, 5*(9), 910-916.

MacLin, O. H., & Webster, M. A. (2001). Influence of adaptation on the perception of distortions in natural images. *Journal of Electronic Imaging, 10*(1), 100-109.

MacPherson, S. (2001). Are faces unique? A comparative investigation of inversion effects in face and object recognition. Unpublished Honours thesis, The Australian National University, Canberra.

Maddess, T., Mccourt, M. E., Blakeslee, B., & Cunningham, R. B. (1988). Factors Governing the Adaptation of Cells in Area-17 of the Cat Visual-Cortex. *Biological Cybernetics, 59*(4-5), 229-236.

Maurer, D., Le Grand, R. L., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6*(6), 255-260.

McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience, 9*(5), 605-610.

McKone, E. (2004). Isolating the Special Component of Face Recognition: Peripheral Identification and a Mooney Face. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30*(1), 181-197.

McKone, E., Aitkin, A., & Edwards, M. (submitted). Categorical and coordinate relations in faces, or is it Fechner's law and face-space instead?
McKone, E., & Boyer, B. (submitted). Four-year olds are sensitive to second-order relational changes in face distinctiveness.
McKone, E., & Grenfell, T. (1999). Orientation invariance in naming rotated objects: individual differences and repetition priming. *Perception and Psychophysics, 61*(8), 1590-1603.
McKone, E., & Kanwisher, N. (in press). Does the human brain process objects of expertise like faces? A review of the evidence. In S. Dehaene, J. R. Duhamel, M. Hauser & G. Rizzolatti (Eds.), *From monkey brain to human brain.* Cambridge, Mass.: MIT Press.
McKone, E., Martini, P., & Nakayama, K. (2001). Categorical perception of face identity in noise isolates configural processing. *Journal of Experimental Psychology: Human Perception and Performance, 27*(3), 573-599.
McKone, E., Martini, P., & Nakayama, K. (2003). Isolating holistic processing in faces (and perhaps objects). In M. Peterson & G. Rhodes (Eds.), *Perception of Faces Objects and Scenes: Analytic and Holistic Processes.* (Vol. 2., pp. 177-211.). Oxford: Oxford University Press.
McNeil, J. E., & Warrington, E. K. (1993). Prosopagnosia: a face-specific disorder. *Quarterly Journal of Experimental Psychology: A, 46*(1), 1-10.
Meadows, J. C. (1974). Anatomical Basis of Prosopagnosia. *Journal of Neurology Neurosurgery and Psychiatry, 37*(5), 489-501.
Mondloch, C. J., Geldart, S., Maurer, D., & Le Grand, R. (2003). Developmental changes in face processing skills. *Journal of Experimental Child Psychology, 86*(1), 67-84.
Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception, 31*, 553-566.
Mondloch, C. J., Lewis, T. L., Budreau, D. R., Maurer, D., Dannemiller, J. L., Stephens, B. R., et al. (1999). Face perception during early infancy. *Psychological Science, 10*(5), 419-422.
Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: a two-process theory of infant face recognition. *Psychological Review, 98*(2), 164-181.
Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Psychology, 9*(5), 555-604.
Murray, J. E., Yong, E., & Rhodes, G. (2000). Revisiting the perception of upside-down faces. *Psychological Science, 11*(6), 492-496.
Nederhouser, M., Mangini, M. C., Biederman, I., & Kazunori, O. (2002). Matching face-like objects is invariant to differences in direction of contrast. Paper presented at the Object Perception And Memory conference, Kansas City, Kansas, USA.
Nicholas, A. K. (1983). *"The Book of the Labrador Retriever"*. Neptune City, NJ: TFH Publications, Inc. Ltd.
O'Toole, A. J., Abdi, H., Deffenbacher, K. A., & Valentin, D. (1995). A perceptual learning theory of the information in faces. In T. Valentine (Ed.), *Cognitive and computational aspects of face recognition* (pp. 159-182.). London: Routledge.
Palermo, R., & Rhodes, G. (2002). The influence of divided attention on holistic face perception. *Cognition, 82*(3), 179-282.

Palmeri, T. J., Wong, A. C.-N., & Gauthier, I. (2004). Computational approaches to the development of perceptual expertise. *Trends in Cognitive Sciences, 8*(8), 378-386.

Parkin, A. J., & Williamson, P. (1986). Patterns of cerebral dominance in wholistic and featural stages of facial processing. In H. Ellis, M. A. Jeeves, F. Newcombe & A. Young (Eds.), *Aspects of face processing* (Vol. 28, pp. 223-227.). Dordrecht: Martinus Nijhoff Publishers.

Parr, L. A., Dove, T., & Hopkins, W. D. (1998). Why faces may be special: evidence of the inversion effect in chimpanzees. *Journal of Cognitive Neuroscience, 10*(5), 615-622.

Pascalis, O., & Bachevalier, J. (1998). Face recognition in primates: a cross-species study. *Behavioural Processes, 43*(1), 87-96.

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science, 296*(5571), 1321-1323.

Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., & Fabregrenet, M. (1995). Mothers Face Recognition by Neonates - a Replication and an Extension. *Infant Behaviour and Development, 18*(1), 79-85.

Pellicano, E., & Rhodes, G. (2003). Holistic processing of faces in preschool children and adults. *Psychological Science, 14*(6), 618-622.

Pellicano, E., Rhodes, G., & Peters, M. (submitted). Are preschooolers sensitive to configural information in faces?

Quinn, P. C., & Eimas, P. D. (1996). Young infants' use of facial information in the categorical differentiation of natural animal species: the effect of inversion. *Infant Behaviour and Development, 19*, 381-384.

Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychological Science, 14*(4), 302-308.

Regan, D., & Hamstra, S. J. (1992). Shape-Discrimination and the Judgment of Perfect Symmetry - Dissociation of Shape from Size. *Vision Research, 32*(10), 1845-1864.

Reicher, G. M. (1969). Perceptual Recognition as a Function of Meaningfulness of Stimulus Material. *Journal of Experimental Psychology, 81*(2), 275-&.

Rhodes, G. (1988). Looking at faces: first-order and second-order features as determinants of facial appearance. *Perception, 17*(1), 43-63.

Rhodes, G. (1993). Configural Coding, Expertise, and the Right-Hemisphere Advantage for Face Recognition. *Brain and Cognition, 22*(1), 19-41.

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition, 47*(1), 25-57.

Rhodes, G., Brennan, S., & Carey, S. (1987). Identification and ratings of caricatures: implications for mental representations of faces. *Cognitive Psychology, 19*(4), 473-497.

Rhodes, G., Byatt, G., Michie, P. T., & Puce, A. (2004). Is the fusiform face area specialized for faces, individuation, or expert individuation? *Journal of Cognitive Neuroscience, 16*(2), 189-203.

Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W. G., & Nakayama, K. (2003). Fitting the mind to the world: Face adaptation and attractiveness aftereffects. *Psychological Science, 14*(6), 558-566.

Rhodes, G., Jeffery, L., Watson, T., E., J., C, W., & Clifford, C. (2004). Orientation-contingent face aftereffects and implications for face coding mechanisms. Current Biology, 14, 2119-2123.

Rhodes, G., & McLean, I. G. (1990). Distinctiveness and expertise effects with homogeneous stimuli: towards a model of configural coding. *Perception, 19*(6), 773-794.

Rhodes, G., Robbins, R., Jaquet, E., McKone, E., Jeffery, L., & Clifford, C. (in press). Adaptation and Face Perception – How Aftereffects Implicate Norm Based Coding of Faces. In G. Rhodes & C. Clifford (Eds.), *Fitting the Mind to the World: Adaptation and aftereffects in high level vision.* Oxford: Oxford University Press.

Rhodes, G., Sumich, A., & Byatt, G. (1999). Are average facial configurations attractive only because of their symmetry? *Psychological Science, 10*(1), 52-58.

Rhodes, G., Tan, S., Brake, S., & Taylor, K. (1989). Expertise and configural coding in face recognition. *British Journal of Psychology, 80*(Pt 3), 313-331.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382-439.

Rossion, B., Dricot, L., Devolder, A., Bodart, J. M., Crommelinck, M., de Gelder, B., et al. (2000). Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience, 12*(5), 793-802.

Rossion, B., Gauthier, I., Goffaux, V., Tarr, M. J., & Crommelinck, M. (2002). Expertise training with novel objects leads to left-lateralized facelike electrophysiological responses. *Psychological Science, 13*(3), 250-257.

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., et al. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: an electrophysiological account of face-specific processes in the human brain. *Neuroreport, 11*(1), 69-74.

Rossion, B., Schiltz, C., & Crommelinck, M. (2003). The functionally defined right occipital and fusiform "face areas" discriminate novel from visually familiar faces. *NeuroImage, 19*(3), 877-883.

Rouse, H., Donnelly, N., Hadwin, J. A., & Brown, T. (2004). Do children with autism perceive second-order relational features? The case of the Thatcher illusion. *Journal of Child Psychology and Psychiatry, 45*, 1-13.

Rouw, R., & de Gelder, B. (2002). Impaired face recognition does not preclude intact whole face perception. *Visual Cognition, 9*(6), 689-718.

Scapinello, K. F., & Yarmey, A. D. (1970). The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. *Psychonomic Science, 21*(6), 329-331.

Schweich, M., & Bruyer, R. (1993). Heterogeneity in the Cognitive Manifestations of Prosopagnosia - the Study of a Group of Single Cases. *Cognitive Neuropsychology, 10*(6), 529-547.

Seitz, K. (2002). Parts and wholes in person recognition: Developmental trends. *Journal of Experimental Child Psychology, 82*(4), 367-381.

Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology, 75*(Pt 2), 221-242.

Simion, F., Cassia, V. M., Turati, C., & Valenza, E. (2001). The origins of face perception: Specific versus non-specific mechanisms. *Infant and Child Development, 10*(1-2), 59-65.

Smith, A. T., & Edgar, G. K. (1994). Antagonistic Comparison of Temporal Frequency Filter Outputs as a Basis for Speed Perception. *Vision Research, 34*(2), 253-265.

Stevenage, S. V. (1998). Which twin are you? A demonstration of induced categorical perception of identical twin faces. *British Journal of Psychology, 89*, 39-57.

Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662.
Subramaniam, S., & Biederman, I. (1997). *Does contrast reversal affect object recognition?* Paper presented at ARVO, Florida, May 1997.
Suzuki, S. (2001). Attention-dependent brief adaptation to contour orientation: a high-level aftereffect for convexity? *Vision Research, 41*(28), 3883-3902.
Suzuki, S. (2003). Attentional selection of overlapped shapes: a study using brief shape aftereffects. *Vision Research, 43*(5), 549-561.
Suzuki, S., & Cavanagh, P. (1998). A shape-contrast effect for briefly presented stimuli. *Journal of Experimental Psychology-Human Perception and Performance, 24*(5), 1315-1341.
Takane, Y., & Sergent, J. (1983). Multidimensional scaling models for reaction times and same-different judgements. *Psychometrika, 48*(3), 393-423.
Tanaka, J. W. (2001). The entry point of face recognition: evidence for face expertise. *Journal of experimental psychology: General, 130*(3), 534-543.
Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science, 12*(1), 43-47.
Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology: A, 46*(2), 225-245.
Tanaka, J. W., & Gauthier, I. (1997). Expertise in object and face recognition. In R. L. Goldstone, D. L. Medin & P. G. Schyns (Eds.), *Mechanisms of Perceptual Learning* (Vol. 36, pp. 83-125.). San Diego, CA.: Academic Press.
Tanaka, J. W., Kay, J. B., Grinnell, E., Stansfield, B., & Szechter, L. (1998). Face recognition in young children: When the whole is greater than the sum of its parts. *Visual Cognition, 5*(4), 479-496.
Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: evidence from a cross-cultural study. *Cognition, 93*(1), B1-B9.
Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory and Cognition, 25*(5), 583-592.
Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise - is the basic level in the eye of the beholder? *Cognitive Psychology, 23*(3), 457-482.
Tarr, M. J. (1995). Rotating objects to recognize them: a case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin & Review, 2*(1), 55-82.
Tarr, M. J. (2003). Visual object recognition: Can a single mechanism suffice? In M. Peterson & G. Rhodes (Eds.), *Perception of Faces Objects and Scenes: Analytic and Holistic Processes*. (Vol. 2., pp. 177-211.). Oxford: Oxford University Press.
Tarr, M. J., & Cheng, Y. D. (2003). Learning to see faces and objects. *Trends in Cognitive Sciences, 7*(1), 23-30.
Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21*(2), 233-282.
Thompson, L. A., Madrid, V., Westbrook, S., & Johnston, V. (2001). Infants attend to second-order relational properties of faces. *Psychonomic Bulletin & Review, 8*(4), 769-777.
Thompson, P. (1980). Margaret Thatcher: a new illusion. *Perception, 9*(4), 483-484.
Tong, F., & Nakayama, K. (1999). Robust representations for faces: evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 1016-1035.

Tong, F., Nakayama, K., Moscovitch, M., Weinrib, O., & Kanwisher, N. (2000). Response properties of the human fusiform face area. *Cognitive Neuropsychology, 17*(1-3), 257-279.

Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron, 21*(4), 753-759.

Tsao, D. Y., Freiwald, W. A., Knutsen, T. A., Mandeville, J. B., & Tootell, R. B. H. (2003). Faces and objects in macaque cerebral cortex. *Nature Neuroscience, 6*(9), 989-995.

Valentine, T. (1988). Upside-down faces: a review of the effect of inversion upon face recognition. *British Journal of Psychology, 79*(Pt 4), 471-491.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology: A, 43*(2), 161-204.

Valentine, T., & Bruce, V. (1986). The effects of distinctiveness in recognising and classifying faces. *Perception, 15*(5), 525-535.

Valentine, T., & Bruce, V. (1988). Mental rotation of faces. *Memory and Cognition, 16*(6), 556-566.

Vuilleumier, P. (2000). Faces call for attention: evidence from patients with visual extinction. *Neuropsychologia, 38*(5), 693-700.

Vuilleumier, P., & Sagiv, N. (2001). Two eyes make a pair: facial organization and perceptual learning reduce visual extinction. *Neuropsychologia, 39*(11), 1144-1149.

Wachsmuth, E., Oram, M. W., & Perrett, D. I. (1994). Recognition of Objects and Their Component Parts - Responses of Single Units in the Temporal Cortex of the Macaque. *Cerebral Cortex, 4*(5), 509-522.

Walton, G. E., & Bower, T. G. R. (1993). Newborns form "prototypes" in less that 1 minute. *Psychological Science, 4*(3), 203-205.

Watanabe, S., Kakigi, R., & Puce, A. (2003). The spatiotemporal dynamics of the face inversion effect: A magneto- and electro-encephalographic study. *Neuroscience, 116*(3), 879-895.

Watson, T., & Clifford, C. W. G. (2003). Pulling faces: an investigation of the face distortion aftereffect. *Perception, 32*, 1109-1116.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature, 428*(6982), 557-561.

Webster, M. A., & MacLin, O. H. (1999). Figural aftereffects in the perception of faces. *Psychonomic Bulletin & Review, 6*(4), 647-653.

Werker, J. F., & Tees, R. C. (1984). Cross-Language Speech-Perception - Evidence for Perceptual Reorganization During the 1st Year of Life. *Infant Behavior & Development, 7*(1), 49-63.

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology, 1*(1), 59-85.

White, M. (2002). Different spatial-relational information is used to recognise faces and emotional expressions. *Perception, 31*(6), 675-682.

Williams, T. M., Fryer, M. L., & Aiken, L. S. (1977). Development of Visual-Pattern Classification in Preschool-Children - Prototypes and Distinctive Features. *13*(6), 577-584.

Wojciulik, E., Kanwisher, N., & Driver, J. (1998). Covert visual attention modulates face-specific activity in the human fusiform gyrus: fMRI study. *Journal of Neurophysiology, 79*(3), 1574-1578.

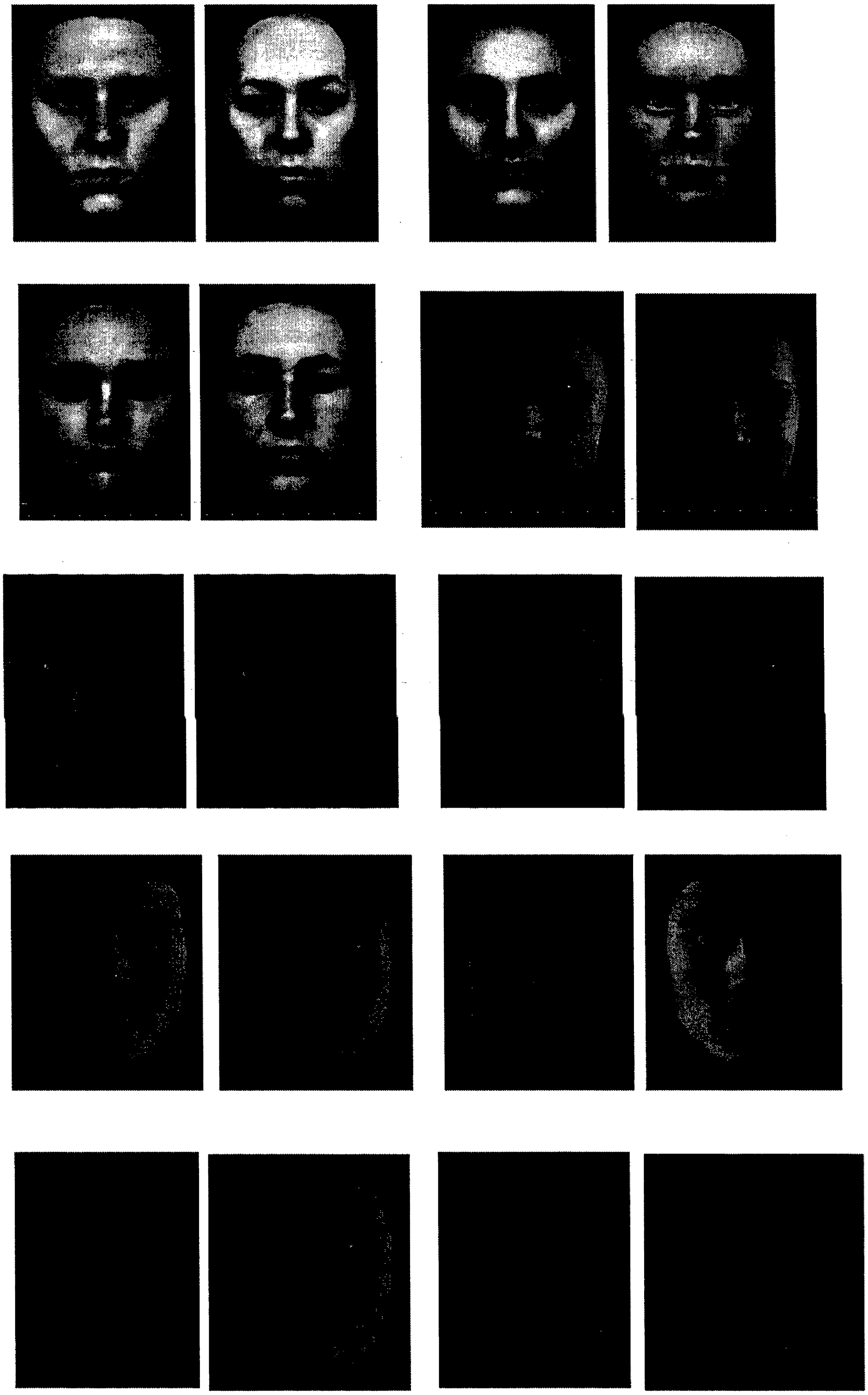
Xu, Y., Liu, J., & Kanwisher, N. (2004). The M170 is selective for faces, not for expertise. *Neuropsychologia.*

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*(1), 141-145.

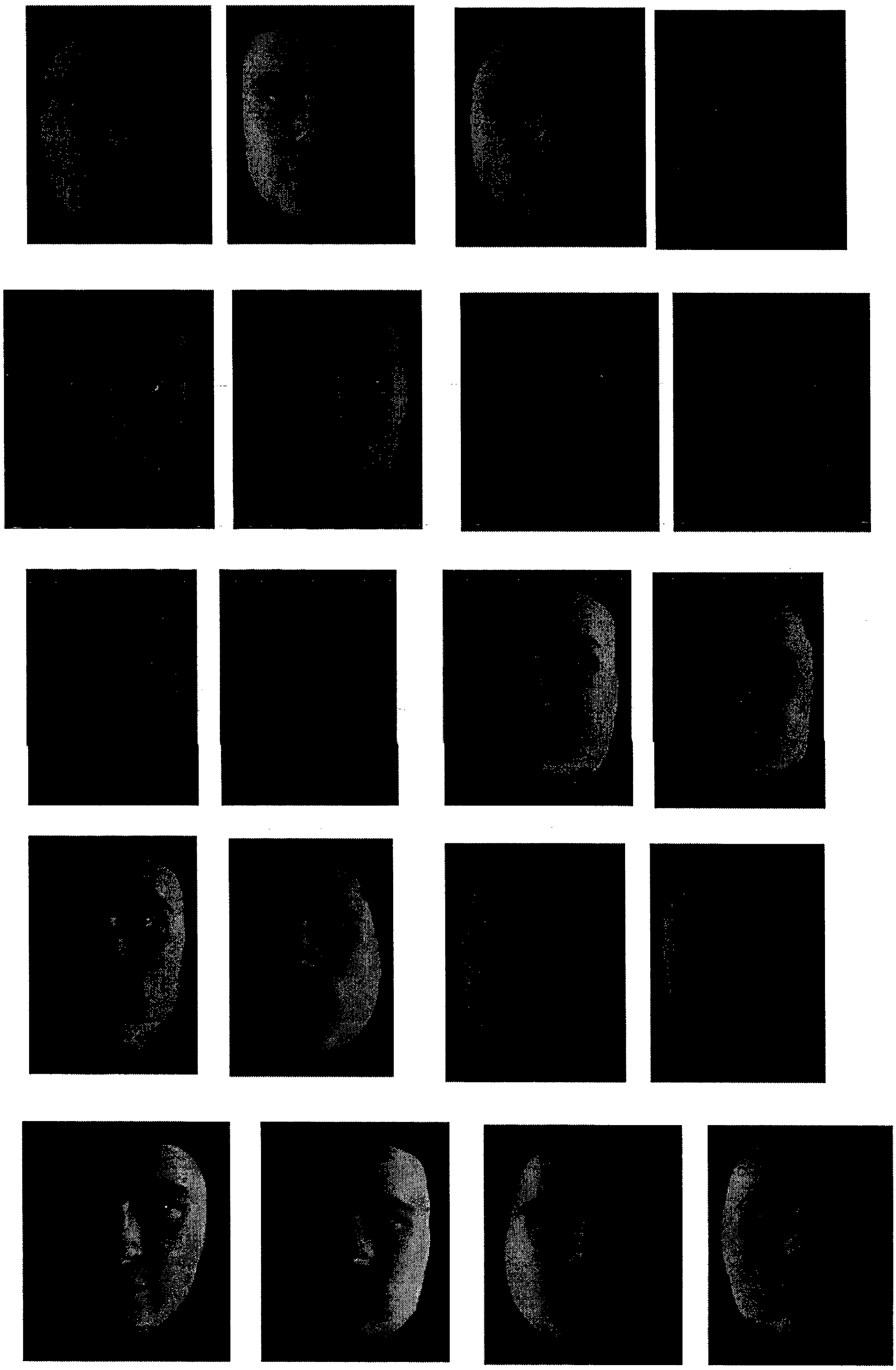
Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*(6), 747-759.

Zhao, L., & Chubb, C. (2001). The size-tuning of the face-distortion after-effect. *Vision Research, 41*(23), 2979-2994.

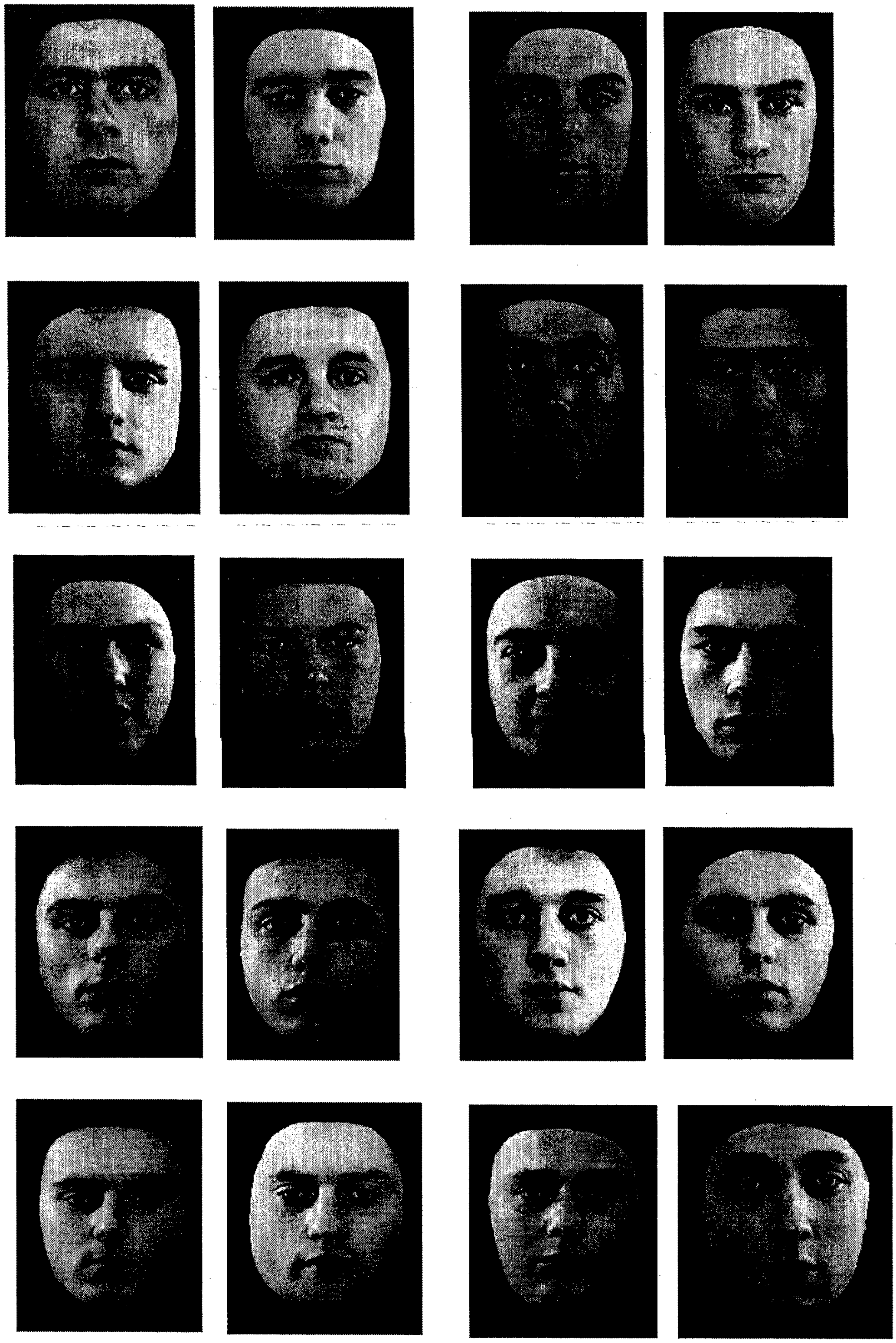
Appendix I: The complete set of face stimuli from Chapters 4 and 5, paired as for Experiments 4 and 7.



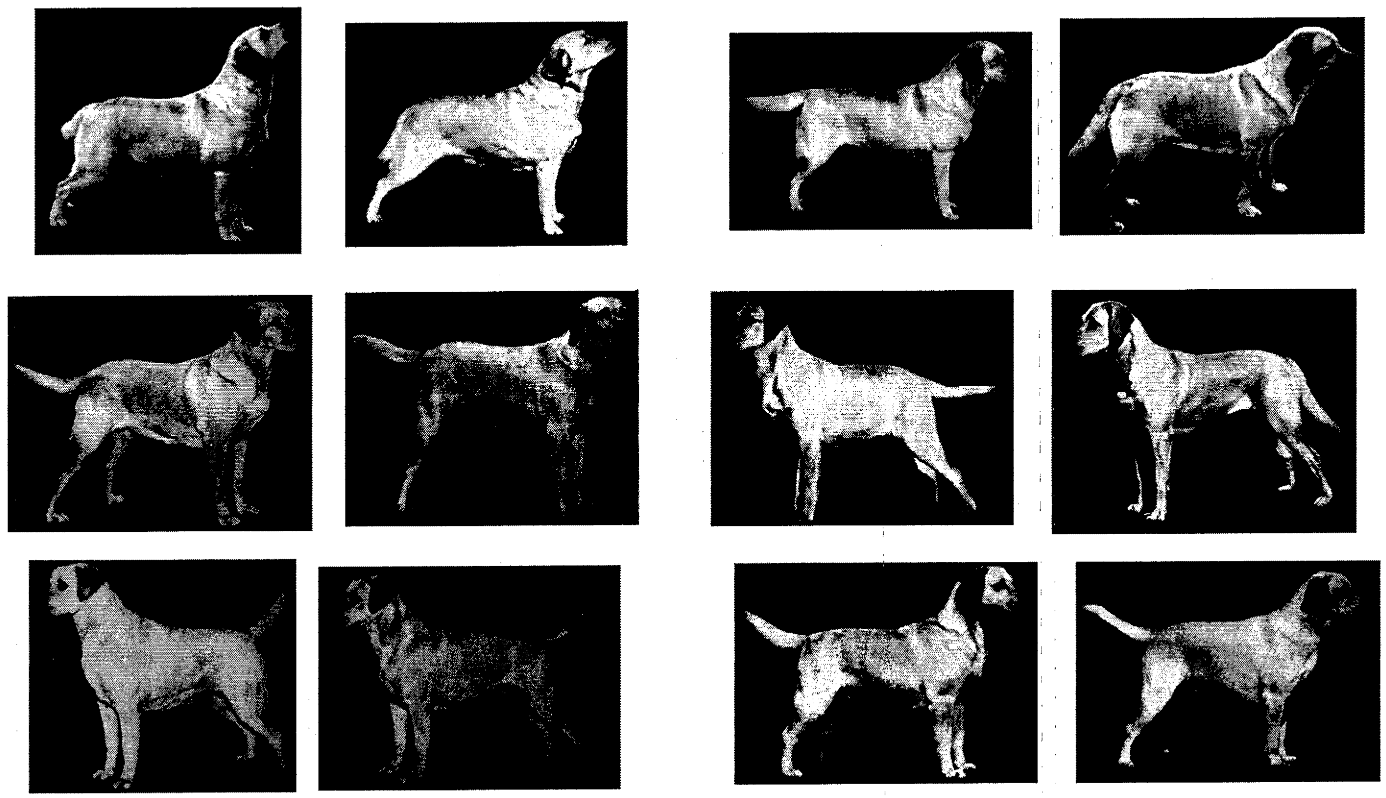
Appendix I: Face stimuli (Chapters 4 and 5)



Appendix I: Face stimuli (Chapters 4 and 5)

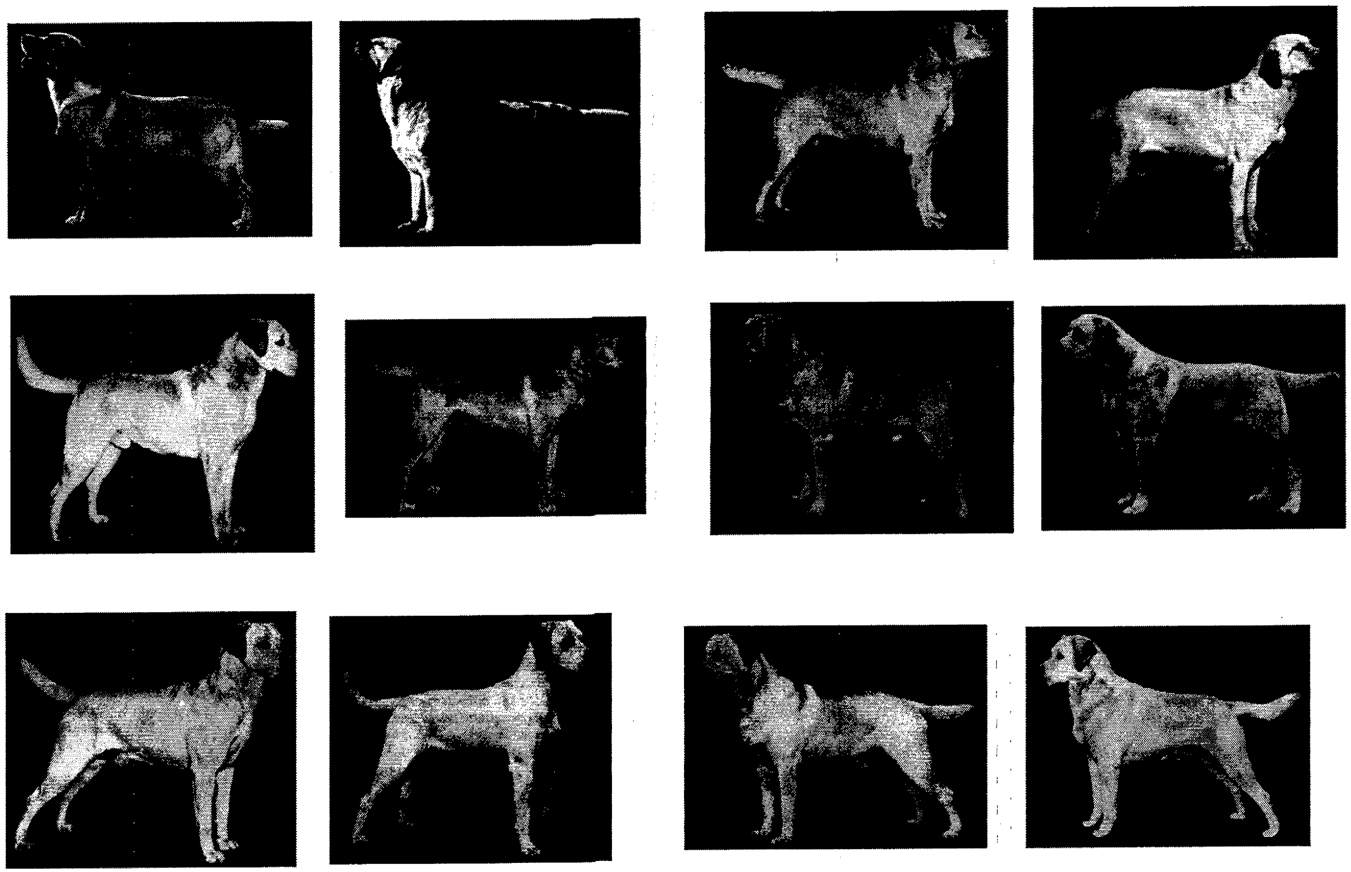


Appendix II: The complete set of dog stimuli from Chapters 4 and 5, paired as for Experiments 4 and 7.

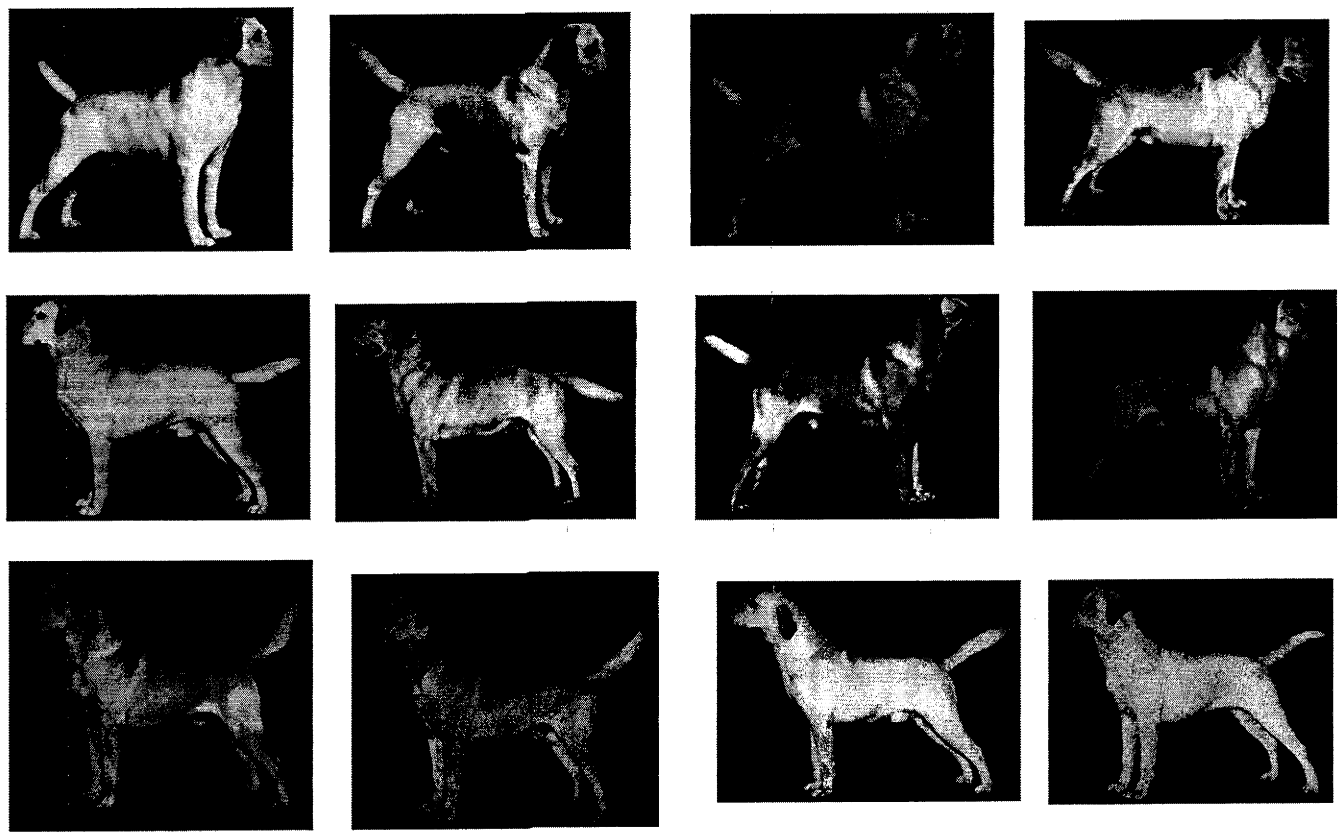


Appendix II: Dog stimuli (Chapters 4 and 5)

Appendix II: Dog stimuli (Chapters 4 and 5)



Appendix II: Dog stimuli (Chapters 4 and 5)



Appendix II: Dog stimuli (Chapters 4 and 5)

Appendix III – details of level of expertise in dog experts

			Dog- shows			Judging			Training	
	Sex	Age	yrs attended	freq/yr	total labs seen (range)	yrs judged	freq/yr	total labs judged	yrs trained	total labs trained
S01	M	76	>37	25	8850-14400	37	25	8850-14400	-	20-30 [b]
S02	M	66	25	20	400-1600	-	-	-	20	4
S03	F	46	42	50	10500	5	6	150-200	-	-
S04	F	59	30	12	1800	25	2	1500	-	-
S05	M	64	32	12	1800	30	3	240-2400	-	-
S06	M	55	23	6	46	15	2	1500-3600	-	-
S07	F	56	23	6	46	16	2	1500-3600	-	-
S08	F	60	34	25	5100-42500	22	3	660-6600	-	-
S09	M	60	35	45	50	7	2.5	70-630	-	-
S10	F	62	8	1	800-960	-	-	-	-	-
S11	M	70	8	1	800-960	-	-	-	-	-
S12	F	41	5 [a]	1	35-50	-	-	-	5	4
S13	F	51	19	20	250	14	3	420-1680	8 [c]	1152-1536 [c]
S14	M	63	22	6	660-5280	18	3	1005-2430	-	-
S15	F	44	28	20	3080-24640	27	2	270-2160	-	-
Mean		58		16.67			4.87			
> 20 yrs experience		59		18.20			5.33			

… table continues next page…

Notes: [a] Refers to guide-dog/puppy meets. [b] Refers to other judges trained rather than dogs. [c] Refers to animal transport business, therefore years in business and number of dogs transported.

	Breeding/showing				Totals (looking at labs)			
	yrs bred	total labs bred	yrs shown	total labs shown	yrs looking at labs	Minimum labs seen	Middle of range (labs seen)	Maximum labs seen
S01	-	-	-	-	37	8870	11635	14400
S02	-	-	-	-	25	400	1100	1800
S03	-	-	-	2	42	10652	10677	10702
S04	-	-	-	-	30	3300	3300	3300
S05	-	-	-	-	32	2040	3120	4200
S06	23	10-230	23	10-230	23	1556	2716	3876
S07	23	10-230	23	10-230	23	1556	2716	3876
S08	-	-	-	-	34	5760	8305	10850
S09	-	-	-	-	10	120	400	680
S10	8	8-32	-	-	8	808	900	992
S11	8	8-32	-	-	8	808	900	992
S12	-	-	-	-	5	35	42.5	50
S13	-	-	-	-	19	1822	2644	3466
S14	-	-	-	-	22	1665	4687.5	7710
S15	-	-	28	56-112	28	3406	15159	26912
Mean					23.07	2853.20	4553.47	6253.73
> 20 yrs experience					29.60	3920.50	6341.55	8762.60

Note: S02 had written a book on training labradors which involved some photographs of dogs. The number of individuals was small (<10) and this is not recorded anywhere in the table.